DOCUMENT RESUME

ED 255 762 CE 041 310

AUTHOR Farley, Joanne; And Others
TITLE Reconceptualization of Vocational Education Program Evaluation.
INSTITUTION Ohio State Univ., Columbus. National Center for Research in Vocational Education.
SPONS AGENCY Office of Vocational and Adult Education (ED), Washington, DC.
PUB DATE 85
CONTRACT 300-83-0016
NOTE 169p.; For a related document, see CE 041 309.
PUB TYPE Information Analyses (070) -- Guides - Non-Classroom Use (055)

EDRS PRICE MF01/PC07 Plus Postage.
DESCRIPTORS Action Research; Educational Research; *Evaluation Criteria; *Evaluation Methods; Inquiry; Models; Postsecondary Education; *Program Evaluation; Research Design; *Research Methodology; *Research Needs; Research Problems; Secondary Education; Trend Analysis; *Vocational Education
IDENTIFIERS *Critical Evaluation; Evaluation Research

ABSTRACT

This document is intended to provide evaluation researchers with information about vocational education as a substantive area to be evaluated, existing methodologies for evaluating vocational education programs, and the theoretical basis for a proposed evaluation method termed critical inquiry. Addressed in the individual chapters of the monograph are the following topics: the need for new directions in vocational education program evaluation, the technological treatment of vocational education evaluation, criticisms of the technological treatment, possible alternative for an evaluation framework in vocational education, theoretical bases for critical theory, critical inquiry as an alternative framework for vocational education evaluation, and suggestions from educational research. A 12-page reference list is appended. (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED255762

RECONCEPTUALIZATION OF VOCATIONAL EDUCATION PROGRAM EVALUATION

Joanne Farley
Floyd L. McKinney
Alan Kohan
Michael Smith
Frank Pratzner

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy

The National Center for Research in Vocational Education
The Ohio State University
1960 Kenny Road
Columbus, Ohio 43210-1090

1985

CE 041 310

THE NATIONAL CENTER MISSION STATEMENT

The National Center for Research in Vocational Education's mission is to increase the ability of diverse agencies, institutions, and organizations to solve educational problems relating to individual career planning, preparation, and progression. The National Center fulfills its mission by:

- Generating knowledge through research
- Developing educational programs and products
- Evaluating individual program needs and outcomes
- Providing information for national planning and policy
- Installing educational programs and products
- Operating information systems and services
- Conducting leadership development and training programs

FUNDING INFORMATION

Project Title:	National Center for Research in Vocational Education, Evaluation Function
Contract Number:	300830016
Project Number:	0510C50010
Act under Which Funds Administered:	Education Amendments of 1976, P.L. 94-482
Source of Contract:	Office of Vocational and Adult Education U.S. Department of Education Washington, DC 20202
Contractor:	The National Center for Research in Vocational Education The Ohio State University Columbus, Ohio 43210-1090
Executive Director:	Robert E. Taylor
Disclaimer:	This publication was prepared pursuant to a contract with the Office of Vocational and Adult Education, U.S. Department of Education. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official U.S. Department of Education position or policy.
Discrimination Prohibited:	Title VI of the Civil Rights Act of 1964 states: "No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance." Title IX of the Education Amendments of 1972 states: "No person in the United States shall on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving Federal financial assistance." Therefore, the National Center for Research in Vocational Education Project, like every program or activity receiving financial assistance from the U.S. Department of Education, must be operated in compliance with these laws.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES vii

FOREWORD xi

EXECUTIVE SUMMARY xi

CHAPTER 1. THE NEED FOR NEW DIRECTIONS IN VOCATIONAL EDUCATION PROGRAM EVALUATION 1

Introduction 1
Federal Initiatives in Vocational Education Program Evaluation 2
Criticisms of Current Practice in Vocational Education Evaluation 4
Conceptual Frameworks in Evaluation 6
Key Components of Evaluative Frameworks 8

CHAPTER 2. A TECHNOLOGICAL TREATMENT OF VOCATIONAL EDUCATION EVALUATION 13

A Scientific Treatment of Vocational Education Evaluation 14
Positivism 16
An Economic View of the Subject Matter 19
Human Capital Theory 20
Evaluation Practice under the Technological Framework 27
Evaluation Criteria 28
Functions of Evaluation under the Technological Framework 29
Objects of Evaluation under the Technological Framework 30
Methods of Evaluation under the Technological Framework 31
Conclusions about the Technological Treatment of Vocational Education Evaluation 35

CHAPTER 3. CRITICISMS OF THE TECHNOLOGICAL TREATMENT OF VOCATIONAL EDUCATION EVALUATION 37

Categorical Inadequacies 37
Subject vs. Object 38
Meaning vs. Reality 38
Subjectivity vs. Objectivity 39

Normative Inadequacy of the Technological Evaluation Framework 40
Value vs. Fact 40
Instrumental vs. Categorical Value Judgements . 41
Lack of Evaluative-Normative Dialogue 43
Inadequate Treatment of the Subject Matter of Vocational Education 44
Vocational Education as Training within the Technological Framework 44
Noneconomic vs. Economic Dimensions 45
Inadequacies in Evaluation Methodology and Practice 46
Overemphasis on Separate and Discrete Outcomes 46
Overemphasis on the Quantifiable 47
Overemphasis on Scientific Explanation 49
Managerial Orientation of the Technological Framework 50
Expert Domination of the Technological Framework 50
Agenda for an Alternative Evaluation Framework . . . 51

CHAPTER 4. POSSIBLE ALTERNATIVES FOR AN EVALUATION FRAMEWORK IN VOCATIONAL EDUCATION 55

Objectivist Evaluation Models 55
The Behavioral Objectives Model 56
The Decision-making Evaluation Model 59
Goal-free Evaluation 63
Subjectivist Evaluation Models 69
Interpretivist Account of Social Inquiry 70
The Case-Study Model. 74
Judicial Evaluation Methodology 81
The Art Criticism Model of Evaluation 85
The Professional Review Model 89
Summary Comments Regarding the Objectivist and Interpretivist Models of Evaluation 90

CHAPTER 5. THEORETICAL BASIS OF CRITICAL THEORY 93

Human Action, Interests, and Knowledge 94
Ideologies and Social Evolution 96
The Ideal Speech Situation 98
The Methodology of Critical Theory 101
Psychoanalysis as a Critical Science 102
Critical Social Analysis 104

CHAPTER 6. CRITICAL INQUIRY AS AN ALTERNATIVE FRAMEWORK FOR VOCATIONAL EDUCATION EVALUATION 107

View and Treatment of the Subject Matter 107
Context for Generating Evaluation Knowledge 109
Framing Evaluation Problems 110
Interpreting Problem Dynamics 110
Overcoming Problems 114
Strengths of the Critical Evaluation Approach . . . 114
Understanding the Past 114
Communication 115
Stakeholder Involvement 115
Continuing Evaluation Effort 115
Constraints on the Conduct of Critical Evaluation . 116

CHAPTER 7. SUGGESTIONS FROM EDUCATIONAL RESEARCH 117

Curriculum Theory: The Behavioral Objectives Approach 117
Job and Task Analysis in Curriculum Construction 121
Competency-based Education 122
Critique of the Behavioral Objectives Approach 123
School-based Learning Theory 126
Bloom's School-based Learning Model 127
Research Findings on School-based Learning. . . 128
Critique of Bloom's School-based Learning Model 136
Effective Schools Research 136
Organization Theory 140
Organization Theory and Design 141
Bureaucracy and Size 142
Technology and Structure 142
Functional vs. Product Structure 142
Environment 142
Goals and Effectiveness 143
Innovation and Change 143
Decision Making 143
Power and Politics 144
Intergroup Relationships 144
Information and Control 145
Top Management Domain 146

REFERENCES 149

LIST OF TABLES AND FIGURES

Table

1 AREAS OF HIGH, MODERATE, AND LOW DEVELOPMENT IN ORGANIZATION THEORY 141

Figure

1. Critical evaluation in vocational education 111

2. Major variables in the theory of school learning 128

3. Dimensions of effective schooling 139

FOREWORD

The last two decades have seen the expenditure of considerable resources to evaluate programs of vocational education. At the same time, there has been intensified criticism leveled at the usefulness of evaluation findings, the criteria used in evaluation efforts, and the methodology employed to evaluate vocational education programs. If program evaluation is to point the direction for program improvement in vocational education, then vocational educators must be willing to reexamine program evaluation and to identify and resolve the crucial problems limiting the effectiveness of evaluation.

Program evaluation is heavily influenced by the arena in which it operates. Accordingly, this manuscript attempts to deal with the philosophical views operating in vocational education programs in addition to proposing a new way of evaluating vocational education programs. As a review of the manuscript will show, these views are complex, often at odds with views within the larger education and community arenas, and seldom discrete as they are operationalized within a school district or even a given school building.

This manuscript is a conceptual piece presenting information about (1) the substantive area (vocational education) to be evaluated, (2) an analysis of vocational education evaluation, and (3) the theoretical basis for the proposed evaluation framework.

The document is intended for use by evaluation researchers in vocational education. Presentation of the theoretical basis for the evaluation framework should also be useful to evaluation researchers working in other substantive fields.

The National Center expresses its appreciation to the following individuals who reviewed the manuscript or provided valuable input: Dr. C. A. Bowers, University of Oregon; Dr. Marjorie Brown, private consultant; Dr. James J. Canelos, The Pennsylvania State University; Dr. Cleo Cherryholmes, Michigan State University; Dr. Donna Coomer, Menominee, Wisconsin, Private Industry Council; Dr. George Copa, University of Minnesota; Dr. John Grasso, West Virginia University; Dr. Ernest House, University of Illinois; Dr. Donald MacKenzie, private consultant; Dr. Mary Malone, Rutgers, The State University; Dr. Michael Patton, University of Minnesota; Dr. Denis C. Phillips, Stanford University; Dr. George Quarles, private consultant; and Dr. Kenneth Sirotnik, University of California, Los Angeles.

The National Center is indebted to the staff members who worked on the study that has resulted in this manuscript. The study was conducted in the Evaluation and Policy Division, Dr. N. L. McCaslin, Associate Director. Dr. Floyd L. McKinney, Senior Research Specialist, served as project director. The project staff included Frank Pratzner, Senior Research Specialist,

and Graduate Research Associates Joanne Farley, Alan Kohan, and Michael Smith. Priscilla Ciulla served as secretary for the project. Editorial review of the paper was provided by Judy Balogh of the National Center's Field Services staff.

Funding for the study was provided by the Office of Vocational and Adult Education, U.S. Department of Education.

Robert E. Taylor
Executive Director
The National Center for Research
in Vocational Education

EXECUTIVE SUMMARY

The focus of this manuscript is the presentation of a conceptual framework within which evaluation of vocational education programs may occur. Initially, the manuscript provides an analytical overview of the nature and function of conceptual frameworks in evaluation. In brief, conceptual or evaluative frameworks provide (1) background assumptions regarding how the object of the evaluation is to be conceptualized, (2) relevant criteria that identify those features or dimensions of the evaluation object which are most relevant in drawing evaluative conclusions, and (3) evaluative criteria to determine the worth or merit of the evaluation object.

This analysis of evaluative frameworks is applied to conventional practice in vocational education program evaluation. The result is the identification of a particular evaluative framework that characterizes much of the current practice in vocational education program evaluation. In the manuscript, this framework is referred to as the technological evaluation framework. The most distinctive features of the technological evaluation framework are its economic interpretation of vocational education and its reliance on scientific methodology in the conduct of program evaluations of vocational education.

Aspects of the technological evaluation framework have been critiqued by researchers, practitioners, and evaluators in vocational education. Moreover, many of the criticisms relate to the technological framework's inadequacy in contributing to local program improvement efforts. These criticisms primarily relate to the following factors: (1) an overemphasis in the technological framework on the economic dimensions of vocational education at the expense of the latter's educational aspects; (2) an overutilization of input-output models of inquiry, which are inadequate in generating an understanding about program processes and context; (3) an overemphasis on generating information to meet accountability rather than improvement concerns; (4) an overemphasis on "expert" knowledge at the expense of practitioners' knowledge about and experience with the program; and (5) insufficient criteria for making evaluative judgments other than those relating to program productivity and efficiency.

The manuscript also provides a review and critique of alternative models of evaluation that might serve to address the inadequacies of the technological evaluation framework. However, the conclusion is that most of the major models currently proposed in the evaluation literature are extensions rather than alternatives to the technological framework, or they generate new problems and inadequacies that would severely limit their usefulness for local program improvement efforts.

A conceptual framework based on critical theory is identified as a potentially fruitful source for evaluation practice

that can inform and guide local vocational education program improvement efforts. Evaluation conducted within the framework of critical theory would generate context-specific knowledge about the particular vocational education program and would be responsive to the needs and interests of groups involved in or affected by the program's operation. A detailed analysis of the basis for critical theory and general suggestions about how it might be applied to vocational education program evaluation are provided. Although the limitations of program evaluation based on critical theory as an alternative evaluation framework are discussed, the conclusion reached is that its strengths and potential contributions to program improvement efforts make it a strong candidate for use in vocational education.

A general overview of other areas of educational and vocational education research is presented in the manuscript. The overview is undertaken from the perspective of how these areas might contribute to local program evaluation efforts. The areas reviewed include curriculum theory, school-based learning theory, effective schools research, and educational organization and administration theory. The conclusion is that review of theory and research in these areas reinforces the adoption of critical theory as a basis for a program evaluation framework in vocational education.

CHAPTER 1

THE NEED FOR NEW DIRECTIONS IN VOCATIONAL EDUCATION EVALUATION

Introduction

Evaluation is not new to vocational education. However, for the first time, the 1976 Vocational Education Amendments introduced more specific evaluation requirements for States receiving Federal funds for the operation of public secondary vocational education programs. Thus, since 1976, there has been a large increase in the amount and types of resources expended on vocational education program evaluation. Yet, during this period, there has been intensified criticism leveled at the content and methodology of much of the program evaluation conducted in vocational education. Much of this criticism can be related to a larger concern that there is a gap in the contributions of evaluation to the practice of providing vocational education to students. Indeed, there is a perception by many that vocational education evaluation has contributed little and has had little impact on efforts to improve programs at the local level.

Where accurate, this charge is a serious one that may affect vocational education's ability to survive. The purpose of this manuscript is to develop a conceptual framework for vocational education evaluation that will address this major problem. More specifically, the manuscript does the following:

- Identifies the ways in which vocational education program evaluation is perceived as inadequate in responding to local practitioners' knowledge and information needs
- Analyzes these inadequacies within the context of broader methodological issues in education program evaluation and of substantive issues revolving around the predominant conceptualization of vocational education's role and function
- Shows that the interpretation of these inadequacies suggests characteristics needed in a reconceptualization of vocational education program evaluation
- Provides a reconceptualization of the methodological and substantive basis for vocational education program evaluation that will bridge the gap between theory and practice

Two major assumptions have guided our inquiry into these issues. The first major assumption made throughout this manuscript is that improvement in the quality and delivery of vocational education is essentially a local affair. As Turnbull (1982) has noted:

> Educational improvement is fundamentally a local enterprise. The people who can bring it about are those who provide educational services--teachers, principals, perhaps local district administrators. (p. 57)

Policies and decisions may be mandated at the Federal and State levels, but unless they are perceived as legitimate by local stakeholders, there is very little probability that they will be actualized in the school. Thus, decisions regarding changes to improve vocational education programs must have the commitment of those individuals whose activities and practices will constitute their operationalization. In this manuscript, an evaluation framework is proposed that encourages local support for legitimate decisions.

The second assumption is that a reasonable and justifiable basis does exist for resolving a major methodological debate that has been ongoing in evaluation theory for some time, that is, the quantitiative (positivist) versus the qualitative (interpretivist) debate. It is recognized that both positivist and interpretivist, or phenomenological, methodologies have something useful to say about the generation of evaluative knowledge; both are also limited and one-sided in their treatment and explanation of human and social phenomena. In this manuscript, a framework is provided that serves as a basis for incorporating into evaluation the positive aspects of these opposing methodologies, while overcoming their respective limitations.

Federal Initiatives in Vocational Education Evaluation

Since the enactment of the Smith-Hughes Act of 1917, State vocational education programs have become increasingly complex and unique. Each State has evolved its own vocational education delivery system funded by Federal, State, and local governments. The Vocational Education Act of 1963 (PL 88-210), the Amendments of 1968 (PL 90-576), and the Amendments of 1976 (PL 94-482) all greatly broadened the scope of vocational education's mandate and target populations. Operating and capital expenditure budgets were enlarged in order to allow vocational education to pursue its legislative mandate. Within this context, pressures have mounted in the last two decades to scrutinize and evaluate vocational education programs. These pressures have come from several sources, including the public's demand for accountability and increased competition among social programs for limited funds.

Prior to 1976, States were relatively free to evaluate their vocational education programs as they saw fit (Cobb and Preskill 1983). As a result, these state-directed evaluations were often sporadic and unsystematic in nature. ABT Associates (1980) summarized the major sources of difficulty identified with State evaluations of vocational education in testimonies before congressional committee hearings. These included the following:

- A lack of adequate planning for vocational education between State and local levels
- Inadequate coordination and planning between State boards of vocational education and higher education agencies, as well as inadequate coordination between either of these entities and State manpower agencies
- A lack of usable and systematic evaluation data on vocational education enrollments, job placements, and program needs

Congress attempted to remedy these sources of difficulty in the 1976 Vocational Education Amendments.

According to Hendrickson (1981), the 1976 amendments envisioned two primary purposes for evaluation: (1) management and improvement of programs and (2) National reporting. With regard to the first priority, Hendrickson noted:

> Congress intended that decisions on which programs to support with Federal funds be influenced by evaluation results, specifically results showing whether or not students from a given program were finding jobs in related fields and whether or not their employers were satisfied with them. (p. 5)

The second priority of the legislation, National reporting, was intended to remedy the lack of an adequate vocational education database. Congress believed that a solid database is crucial to good planning and program improvement at the State and local levels (ABT Associates 1980). Specifically, the legislation established the National Occupational Information Coordinating Committee (NOICC) and, in each State, a State Occupational Information Coordinating Committee (SOICC). The NOICC and SOICC occupational information systems were intended to meet the common occupational information needs of vocational education programs and employment and training programs at the National, State, and local levels. To assist in the collection and reporting of data, a National data collection system, the Vocational Education Data System (VEDS), was established. VEDS has experienced a rather turbulent history and the system was not operational during the 1983-84 school year.

Most recently, the evaluation of local programs of vocational education has been primarily driven by the requirements of the Vocational Education Amendments of 1976 (P.L. 94-482). As long as States received Federal funds that were subsequently targeted to local programs, States complied with the regulations established in the 1976 act. Worthington (1984) noted that the federal-to-state funding ratio was 1 to 10.9 in 1982. Nevertheless, no State has declined Federal funds in exchange for sole State control of vocational education.

The Carl D. Perkins Vocational Education Act of 1984 also designates specific measures of effectiveness for vocational education program evaluation: the occupations to be trained for, which should reflect a realistic assessment of the labor market needs of the State; the levels of skills to be achieved in particular occupations, which should reflect the hiring needs of employers; and the basic employment competencies to be used in performance outcomes, which should also reflect the hiring needs of employers. The requirement for program evaluations by the State once each 5 years continues in the 1984 act.

Criticisms of Current Practice in Vocational Education Evaluation

As indicated in the last section, an increasing amount of time and resources has been expended on vocational education program evaluation since 1976. However, a large number of vocational education researchers and practitioners would nevertheless agree that the following quote is as apt a description today as it was in 1976:

> The literature describing the evaluation of vocational education programs is discouraging: it yields little useful information for vocational educators. The research designs have used analytic procedures requiring simple quantitative input and have failed to encompass many important educational issues. Evaluations have used research methods that are incompatible with the complexity of the learning, teaching, and administrative situations. (Committee on Vocational Education Research and Development 1976, p. 107)

We suggest that many of these and other inadequacies are associated with the operation of a particular conceptual framework underlying much of current vocational education evaluation practice. In chapter 2, this framework is called the "technological" evaluation framework and discussed at some length. Here we will simply assert that it does in fact operate in vocational education program evaluation and has been the focus of much criticism. The criticisms are both theoretical and practical in nature.

Moreover, these criticisms suggest that the technological framework does not and cannot have much relevance for program improvement efforts. In chapter 3 these criticisms are presented in detail. Here, a few of the major inadequacies are summarized with the technological evaluation framework, which indicate a need for new directions in vocational education evaluation.

Current evaluation practice often treats vocational education as basically an economic process that is expected to yield outcomes of an economic nature. This economic treatment frequently imposes on vocational education a simple "input-output" industrial production model. Critics claim that the input-output model is a far too simplistic treatment of what is actually a complex, multifaceted phenomena. Such oversimplification, it is argued, yields information that is not particularly revealing about sources of problems and deficiencies in the delivery of vocational education services. A second criticism often voiced is that evaluations based on the industrial production model reinforce both the focus on vocational education's economic consequences and goals at the expense of other non-economic goals and outcomes. They argue that even though Federal mandates require attention to labor market outcomes, the latter is often not consistent with the values and objectives of local practitioners. Thus, they claim, evaluations should reflect the fact that vocational education serves important social and educational roles that are at least as important as its economic ones.

Another major inadequacy identified in current evaluation practice is its reliance on the empirical-analytic or "positivist" model of scientific inquiry. Adherents of the positivist model of inquiry claim or assume it provides the means for generating knowledge which is as "objective" or nonbiased as possible. Thus, on the positivist account, evaluation deals only with the realm of facts while leaving the activity of valuation to other disciplines and groups of individuals. Critics claim that the positivist model is itself theory laden and normatively bound by prescriptive assumptions. They argue that findings based on the positivist model can only be interpreted by reference to its basic assumptions about human and social phenomena. These in turn imply certain valuative stances toward human needs and the circumstances of human existence. Thus, the dichotomization of fact and value is not only erroneous, but it is also an undesirable characteristic of inquiry aimed at yielding knowledge for valuing human purposes and activities.

The positivist view of scientific inquiry is also perceived by many to be categorically inadequate for the explanation and understanding of human beings. Although perhaps relevant to the study of inanimate objects, categories and concepts such as causality, prediction, generalization, and behavior are seen as inappropriate for application to human beings and their interaction. Evaluations that use inadequate categories of meaning usually yield a great deal of information about statistical correlations

but very little real understanding of the dynamics underlying practical problems in local contexts.

The technological framework is often viewed as inadequate because it is "management oriented" rather than practitioner oriented. That is to say that evaluation activities carried out under this framework usually focus on performance indicators that are expected to yield information relevant to policy and resource allocation decisions. Given the current administrative structure of vocational education, these are management issues and not the issues confronted by practitioners in the classroom or school. Thus, evaluation findings are often irrelevant to practitioners and local stakeholders who are concerned with attempts to improve the quality of vocational education.

A final criticism is that evaluation under the current framework is expert dominated. Vocational education evaluation is often conducted by outside parties with credentials in evaluation methodology. Such persons conduct the evaluation (i.e., design the study, collect and analyze the data, and write reports) with little or no participation from those individuals closest to the program. Without such participation, it is difficult to say how expert-controlled evaluations can be validated or how findings from such evaluations may be useful in local contexts.

As noted before, the inadequacies identified with current evaluation practice will be noted in detail in chapter 3. However, the discussion here indicates that these perceived inadequacies are serious enough to justify taking a new direction in vocational education program evaluation. Were this issue to be merely a matter of technique, there would be no need for this manuscript. Instead, it is suggested that the identified inadequacies are more fundamentally related to the operation of a particular conceptual framework that prescribes how evaluative knowledge is to be acquired and, thus, what aspects of reality are to be evaluated. In this manuscript, the attempt is to make this framework explicit and to explicate an alternative framework that is implied in the criticisms of the former. First, the interpretation of the role of conceptual frameworks in evaluation is briefly discussed.

Conceptual Frameworks in Evaluation

Imre Lakatos (1970) noted that "for centuries knowledge meant proven knowledge--proven either by the power of the intellect or by the evidence of the senses" (p. 91). However, since the replacement of Newtonion physics by Einstein's theoretical success, "now very few philosophers or scientists still think that scientific knowlege is, or can be, proven knowledge" (p. 91). This acknowledgement has thrown open the whole question of the credibility of scientific knowledge. Having dislodged the

twin pillars of empiricism and formal logic, philosophers of science are taking a closer look at what it is and what it historically has been that scientists do when they successfully practice scientific inquiry and generate scientific knowledge.

Although disagreements abound among scholars regarding the accurate and appropriate description of scientific inquiry (refer to Lakatos and Musgrave 1970), there is (at least) one point that has received widespread, general agreement. This point has to do with the fact that scientific inquiry is driven and guided by some set of concepts, assumptions, and meanings that guide the scientist's interpretation of the world and also his or her methodological treatment of that phenomena. These systems have been given various names and explanations--such as Kuhn's (1970) "paradigms," Lakatos's (1970) "research programmes," Laudan's (1970) "research traditions," Toulmin's (1961) "ideals of natural order." Yet, explicit in all accounts of such systems is the rejection of the classical empiricist tenet that scientific knowledge is obtained by the imprint of nature on passive and neutral minds. Instead, most of current philosophy recognizes that the scientist's manner of selecting phenomena as relevant and making sense of it is guided by some interpretive framework or "lens" through which the scientist looks upon the world.

In this respect at least, evaluation is no different from other types of scientific pursuit. To the extent that evaluation is an activity aimed at the generation of knowledge (i.e., evaluative knowledge) about some phenomena or states of affairs in the world, it too is driven and guided by interpretive schemas or frameworks. Thus, evaluation, like scientific research, is only conducted in relation to some subject matter. Evaluation may be, as Scriven (1980) claimed, "a subject in its own right" (p. i). However, the fact remains that evaluation is intended to yield knowledge about some substantive entity, for example, an economic policy, a manpower program, or an educational innovation. Just as evaluative results may affect the design, implementation, and history of a program or policy, the substantive entity will also interact with evaluation by providing its subject matter and substantive content. In short, the treatment of this subject matter will direct the evaluation by providing it with a theoretical and conceptual context within which evaluative questions and criteria are identified. It is suggested that the conceptual context within which evaluation is conducted and the methodological context within which evaluative knowledge is acquired together constitute an "evaluative framework."

Briefly stated, it is argued in this section that evaluative frameworks, if made explicit, serve the following interrelated functions:

- They serve to provide the evaluator with background assumptions regarding how the subject matter of the

evaluation is to be treated or interpreted, for example, some form of economic treatment as opposed to a psychological treatment.

- They provide criteria of relevance regarding what dimensions, aspects, and components of the subject matter to be evaluated are necessary to account for or warrant conclusions reached by the evaluation.

- They provide evaluative criteria through which the conclusions reached by the evaluation can be interpreted qua evaluative conclusions (e.g., the evaluative conclusion that a policy is or is not desirable).

Following is a discussion of these functions as they related to key components of evaluative frameworks.

Key Components of Evaluative Frameworks

Background assumptions. Evaluative frameworks contain background assumptions and theories that, taken collectively, provide the conceptual context within which a subject matter is defined and its treatment guided. As Gouldner (1982) pointed out, such assumptions are "'silent partners' in the theoretical enterprise" which "from beginning to end, influence a theory's formulation and the research to which it leads" (p. 324). These background assumptions serve a similar role to that of "generative metaphors." As House (1983) explained:

> Schon maintains that we are guided in our thinking about social policy by pervasive, tacit images that he calls generative metaphors, in which one frame of reference is carried over to another situation. . . . Naming and framing proceed by generative metaphor. . . . The "deep" metaphor accounts for why some elements are included in the story while others are not, some assumptions are taken to be true in spite of disconfirming evidence, and some recommendations seem obvious. (pp. 8-9)

Background assumptions are also like Stephen Toulmin's (1961) notion of "ideals of natural order." Toulmin claimed that the explanatory power of science (in whatever form it takes) is generated by ideals of natural order. These consist of our "prior expectations which in turn reflect our ideas about the order of Nature" (p. 56). Ideals of natural order are accepted as self-explanatory. They specify the way in which things behave of their own nature if left to themselves. It is because of ideals of natural order that scientific explanations "make sense" and "hang together" to make events in nature intelligible. They are never shown (or expected to be shown) to be true or false;

rather, over time, "they 'take us further (or less far)', and are theoretically more or less 'fruitful'" (p. 57).

When explicated, the background assumptions of an evaluative framework will involve some notion of human welfare. Charles Taylor (1974) made this point in the following quote:

> For a given framework is linked to a given conception of the schedule of human needs, wants and purposes, such that, if the schedule turns out to have been mistaken in some significant way, the framework itself cannot be maintained. This is for the fairly obvious reason that human needs, wants and purposes have an important bearing on the way people act, and that therefore one has to have a notion of the schedule which is not too wildly inaccurate if one is to establish the framework for any accurate science of human behaviors. . . . A conception of human needs thus enters into a given political theory, and cannot be considered something extraneous which we later add to the framework to yield a set of value-judgments. (p. 31)

Taylor is arguing that to the extent a given explanatory framework is involved with human welfare and to the extent that it is nontrivial. the framework will have to make some assumptions regarding what it is (or is partially) that constitutes human needs. Moreover, these assumptions cannot be "tacked on" or deleted from the explanatory framework. Rather, like ideals of natural order and generative metaphors, the schedule or conception of human welfare is the core by way of which explanations are developed and elaborated upon.

Finally, the background theories or assumptions of a particular framework will demarcate the appropriate and relevant domain or context within which a particular subject matter is to be located. This context will then suggest how that subject matter is to be substantively and methodologically treated. It does so by providing substantive and methodological criteria of relevance.

Criteria of relevance. A framework's criteria serve to point out relevant features or characteristics that must be considered in evaluating a particular subject matter. Thus, they inform us as to what counts as relevant research and evaluation questions and what type(s) of knowledge are needed to establish the "facts" of the matter. In a similar argument, Taylor (1974) claimed that theories contain "crucial dimensions of variation." According to Taylor, the latter serve to

> delineate the relevant features in the different dimensions and their relation so that we have some idea of what can be the cause of what, or how character affects political process, or social structure affects

> character, or economic relations affect social structure . . . vice versa. . . . Before we have made some at least tentative steps in this direction we don't even have an idea where to look for our explanations; we don't know which fact to gather. (p. 20)

Besides highlighting the relevant dimensions of a phenomenon, criteria of relevance serve to structure the relations between the facts or dimensions. Thus, in some frameworks, the structure of relations between dimensions of the subject matter may be determinate and linear; in others, it may be indeterminate and dialectical. The important point here is that an evaluative framework provides criteria of relevance that not only guide our treatment of a subject matter but also guide the content and structure of the knowledge generated through the filter of its background assumptions and theories.

Evaluative criteria. In providing background assumptions regarding human needs or welfare and related criteria of relevance, an evaluative framework also provides evaluative criteria by which certain states of affairs, policies or programs within the domain of interest can be adjudged desirable or undesirable, acceptable or unacceptable, legitimate or illigitimate. Thus, an evaluative framework's background assumptions and criteria of relevance imply what Phillips (1980) has called "linking premises," which relate value judgments to empirical explanations. To quote from Taylor (1974) again:

> Thus the framework does secrete a certain value position, albeit one that can be overridden. In general we can see this arising in the following way: the framework gives us as it were the geography of the range of phenomena in question, it tells us how they can vary, what are the major dimensions of variation. But since we are dealing with matters which are of great importance to human beings, a given map will have, as it were, its own built-in value-slope. This is to say, a given dimension of variation will usually determine for itself how we are to judge of good and bad, because of its relation to obvious human wants and needs. (p. 29)

Moreover, in demonstrating that a state of affairs or social program fulfills some schedule of human needs or welfare, a framework is also demonstrating that, prima facie, that situation or program is desirable (i.e., that it is "good").

> Many educational policies involved reference to entities commonly regarded as intrinsic goods. Some of these policies take the form "schools should ensure that pupils are X" where X can be replaced by "honest", "chaste", "intellectually curious", "respectful", and so on. How can any of these policies be justified? As soon as the

> question is raised, its oddity becomes apparent. For the policy of practicing an intrinsic good requires no justification. If X is regarded as an intrinsic good, by definition it is good or of value in itself, and the practice of it is, manifestly, justified. (Phillips 1971, p. 55)

When a claim is made that a policy or program is "good," rejection of the claim for merit or worth of the program can only occur validly through the provision or setting forth of "overriding" or "undermining" arguments. Taylor (1974) explained this about an overriding objection: "It may be that the valuation is accepted, but that its verdict for our actual choices is overriden, as it were, by other more important valuations" (p. 28). In other words, we may accept the empirical explanation or claim that the policy in fact fulfills some set of human needs and is, thus, ceteris paribus desirable. However, we may then reject the evaluative conclusion that the policy is desirable because of other more important conclusions (e.g., the policy results in consequences that reduce human welfare more than they increase it). In contrast, an undermining argument or objection is one that "undermines the valuation itself, seeks to deprive the putative good of its status" (ibid.). We deny that the program in question has the properties by which it is judged good (e.g., program X does not in fact fulfill the set of human needs that is claimed for it).

This characterization of an evaluative framework will be used to organize and focus discussion throughout the remainder of this document. Thus, in chapter 2, this account is applied to an analysis of an existing body of belief and knowledge in vocational education. This framework is referred to as the technological evaluation framework. In chapter 3, a detailed critique, concluding that the technological framework has a number of inadequacies that severely limit its appropriateness as an evaluative framework is provided. Moreover, an "agenda" of issues, which any adequate alternative to the technological evaluative framework will be expected to address, is developed. Although the technological framework is the conventional and predominant context within which vocational education evaluation is conducted, there is nevertheless a good deal of diversity in the design features of specific evaluation models. Thus, in chapter 4, we rely with minor modifications on House's (1980) distinction between evaluation models that rest on objectivist epistemological grounds and those that rest on subjectivist epistemological grounds. The former include the systems analysis model, the behavioral objectives model, the decision-making model, and the goal-free model of evaluation. Subjectivist Models include the art criticism model, the peer review or accreditation model, the quasi-legal or adversary model, and the case-study or transactional approach to evaluation. Besides providing an overview of these models, chapter 4 also provides a critique of each in terms

of how it addresses the agenda of issues developed in chapter 3. Whereas the critiques of the models serve to expand the agenda of issues, they also indicate that none of the major evaluation models can adequately address the issues included in the agenda. Thus, a different source of an alternative framework for vocational education must be sought. In chapter 5, what is thought to be a promising area in which to look for just such an alternative is identified. This is the area of social inquiry called critical theory. Chapter 5 then attempts to explicate the theoretical basis of critical inquiry. Finally, chapter 6 explores some of the implications of reconstructing critical inquiry as an evaluative framework. Included in this discussion is consideration of the manner in which critical evaluation inquiry may tackle the agenda of issues generated in earlier chapters. Chapter 6 will also explore some of the major constraints on conducting evaluations from within a critical framework.

CHAPTER 2

A TECHNOLOGICAL TREATMENT OF VOCATIONAL EDUCATION EVALUATION

In this chapter, we present an analysis of what we consider to be the predominant evaluative framework currently operating in vocational education. It should be reemphasized at this point that the analysis represents a reconstruction of the evaluative framework (i.e., we are making explicit what is often implicit and tacit). Thus, no suggestion is being made that all evaluators and vocational educators would agree with this characterization of the framework. However, it is hoped that, when fully explicated, the logic of the framework pushes one to certain conclusions and valuations.

The focus of this chapter is on evaluation in vocational education. However, the technological framework discussed here is operative in vocational education research, education research and evaluation, and other broad areas of social research. Although some of these social areas have begun to examine and to alter their stance concerning appropriate frameworks for conducting research and evaluation activities, most have experienced a period of time when the technological framework was (or remains) the dominant framework.

As discussed in chapter 1, an evaluative framework contains a set of background assumptions that guide definition of the subject matter for evaluation and the way in which one attempts to obtain evaluative knowledge about that subject matter. Thus, these background assumptions shape and inform both the content and methodology of an evaluation. In this section, how the background assumptions contained in the predominant evaluative framework are drawn from theories or conceptual frameworks operating in disciplines other than education or vocational education are discussed. Collectively, they constitute what we describe as a "technological" treatment of vocational education evaluation.

As Bowers (1982) has noted, a number of cultures have "evolved elaborate and sophisticated techniques" for dealing with the conditions of their environment. Yet this alone does not warrant characterizing them as technological cultures. Instead, technicism refers to "a direct relationship between a [unique] pattern of thought and technology" (p. 530). In the West, this relationship evolved out of the confrontation between Enlightenment principles of reason and individual autonomy on the one hand and Christian and feudal principles of faith and subservience to an existing order on the other. An important aspect of this evolution has to do with the particular meanings that the "economic" and the "scientific" have come to hold in the West. It is not our purpose to conduct a full-blown analysis of the evolution of Western technological consciousness. However, it is suggested

that Western technicism is reflected in the operation of a technological evaluation framework in vocational education (at least in the United States). In stating that the predominant evaluation framework constitutes a technological treatment of vocational education, we are saying that it constitutes a particular interpretation of the scientific treatment of an economic subject matter. In short, it is suggested that the content of evaluations conducted within vocational education is usually guided by reference to human capital theory and scientific management concepts. The methodology employed is usually informed by the empirical-analytic or positivist model of scientific inquiry. In the following sections, the background assumptions underlying these distinctive notions of the economic and the scientific are set forth.

A Scientific Treatment of Vocational Education Evaluation

The great gains made in the physical sciences during the Enlightenment gave momentum to the idea that science as opposed to religion or faith was a more reliable foundation for acquiring knowledge about the world. In the 19th century, thinkers such as Comte, Bentham, and J. S. Mill began advocating the application of scientific method to the study of man and society. Interpretations regarding what constitutes science have differed through the years. However, during the late 19th and early 20th century, basically one interpretation began to emerge in those disciplines now called the social sciences (e.g., psychology, sociology, political science, and history). Educational research was not isolated from the scientific movement, as is illustrated by the following quote:

> When the scientific approach is applied to the study of educational problems, educational research is the result. Educational research is the way in which one acquires dependable and useful information about the educative process. Travers defines educational research as "an activity directed toward the development of an organized body of scientific knowledge about the events with which educators are concerned." Its goal is to discover general principles or interpretations of behavior that can be used to explain, predict, and control events in educational situations--in other words, scientific theory. (Ary and Razavieh 1979, p. 21)

During the 1960s when the West was experiencing serious sociopolitical and economic unrest, social scientists became (or, at any rate, were made) aware of the need to move away from "pure" research toward research relevant to the problems of the times. The new watchword of "applied science" came into vogue as disciplines changed focus and even subdivided into new areas of social inquiry. In the public policy domain, the issues of

equity and public accountability became key priorities. It was within this context that evaluation emerged as a separate and distinct discipline. However, in its separation from the research communities, evaluation retained the image of itself as a "scientific" discipline. Baker (1979) said this of educational evaluators:

> People assumed they could apply their research rationality to evaluation problems. . . . It was assumed that evaluation shared the basic precepts of a good science--that evaluation was independent the way research is supposed to be independent; that evauation was orderly; that expertise was required for its conduct; and that, by virtue of training, some people ought to be better at it than others. Supporting beliefs honored the value of measurement, attractability, discovery of causal relationships, suspension of disbelief for questionable human data sources and, of course, the idea of design and control. All these research ideas were transferred to the evaluation framework. (p. 1)

As these quotes indicate, agreement on the need for scientific methods has usually included agreement on what constitutes scientific method. The interpretation placed on the latter has frequently been called the positivist interpretation of scientific inquiry.* There are available a number of in-depth analyses of the positivisitic account of science [refer, for example, to

---

*It should be emphasized that what we are reconstructing here is an interpretation regarding scientific inquiry and methodology. Moreover, we are describing an interpretation that evolved within those fields of inquiry now called the social sciences. We are not suggesting that this in any way accurately reflects what scientists in the "physical" sciences do.

We also agree with Phillips (1983) who said that the equation of positivism in the social sciences with the position taken by logical positivists in the philosophy of science is mistaken both with regard to history and to a reading of the latter's program. Although the logical positivists perhaps made more sense of scientific inquiry than did or have the positivists in the social sciences, they were not the originators of the basic tenets of positivism. Without attempting an intellectual history, we will simply suggest that the hallmarks of positivistic inquiry (i.e., objectification, generalization) emerged with the advances of Newtonian physics in the 17th century. Their application to social and human affairs is reflected in thinkers such as Hobbes, Voltaire, Locke, and more recently Comte, John S. Mill, and Durkheim--some of whom preceded by some years the members of the logical positivist movement in the philosophy of science.

Patton (1975), Phillips (1983), and Smith (1983)]. Thus, only some of the more distinguishing characteristics of positivist social inquiry are pointed out here.

Positivism

The point of positivism "is to construct an objective, empirical and systematic foundation for knowledge" (Held 1980, p. 164). In attempting to achieve this aim, the positivist account makes assumptions about the nature of reality, the appropriate methods for acquiring knowledge, and what it is that constitutes acceptable scientific knowledge. We will discuss these briefly.

The positivist view is that reality is atomistic (i.e., composed of separate and distinct particles) in terms of its basic composition. These particles assume their forms due to uniformities and regularities in their interrelations. The uniform and regular interrelations are, in turn, the effects of causal relations that are maintained between and among particles. Moreover, events in reality are outcomes or products of the causal dynamics of two or more particles or sets of particles. The nature of reality, then, is contingent upon or determined by the recurring interplay of cause and effect. Finally, positivism holds that reality is objective insofar as it exists externally to and independently from human consciousness. Taylor (1975) contrasted this view with the older Aristotelian views of reality.

> The new notion of objectivity rejected the recourse to final causes, it was mechanisitic in the sense of relying on efficient causation only. Connected with this it was atomistic, in that it accounted for change in complex things not by gestalt or holisitic properties, but rather by efficient causal relations among constituents. It tended toward homogeneity in that seemingly qualitatively distinct things were to be explained as alternative constructions out of the same basic constituents or basic principles. . . . Thus this science was mechanisitic, atomistic homogenizing, and of course saw the shape of things as contingent. (p. 9)

Metaphorically speaking, this view of reality is analagous to the way in which one thinks of the operation of a "well-oiled" machine. A machine is put together by the assembly of a number of separate and distinct components or parts. These component parts are interrelated in such a way that their systematic functioning results in the machine's effecting some consequence either in the form of a product or in the form of some operation such as seeding a field. The production of this consequence is not contingent upon any subjective states that may be attributable to the components; indeed, ordinarily we do not think of machine parts as having subjective states of mind. Rather the

consequence can be explained through the application of general principles of dynamics that stand in external relation to any individual part of the machine. Finally, the machine and the consequences or effects may be replicated at any time or place by assembling the same or similar component parts in a similar manner. This points up the static view positivists hold of reality (i.e., the dynamics of the particles and the particles of which reality is comprised do not alter over time).

As Held (1980) noted, after the Enlightenment, modern science "understood the world as a scientific universe which could be comprehended only by itself" (p. 160). Moreover, knowledge on the positivist view has come to be equated with scientific knowledge. Over time, use of the scientific method results in the accumulation of knowledge. This conception of the cumulative nature of scientific knowledge rests on the assumption that the world can be described and explained in terms of lawlike generalizations that specify regular and recurring relations between entities. Although these lawlike generalizations are not themselves subject to direct comparison with empirical reality, they can be used to deduct or predict statements describing particular instances of a general causal relationship. These particular statements are hypotheses. Hypotheses can be operationalized and given empirical import. They can then be tested against empirical data. This emphasis on testability is to ensure the "objectivity" of what is accepted at any point in time as scientific knowledge. Scientific knowledge must be independent of the context within which it is developed and, thus, independent of any one person or group of persons who espouse it. Demanding that theoretically derived hypotheses be testable ensures that tests can be replicated and repeated across varying time periods and across diverse situational contexts. To the extent that test results do not disprove a hypothesis, the hypothesis and the deductive-nomological theory from which it was derived can be considered confirmed (at least for the time being). Finally, it is assumed that over time confirmed theories will be subsumed under broader and more general theories, which in turn can be used to predict ever-increasing numbers of behavior occurrences. Thus, in the positivist model, generalizability, testability, and objectivity are the hallmarks of scientific method.

Besides positing assumptions about the nature of reality and knowledge, positivism also holds to a particular view of meaning and language. A long and much held belief by pre-Enlightenment thinkers was that the order of things was inherently meaningful. Meaning was conceived to be the embodiment of purpose (either spiritual or philosophical) as expressed or reflected in the shape and substance of natural and social reality. Language was thought of as one more mode of expression through which life was provided with meaning. Thus, meaning was considered to be present whether or not humans understood it fully or even were conscious of it. In contrast, positivism assumes that meaning is

a subjective category imposed by humans on reality. As such, it is not an inherent quality of reality. In this view, language is the vehicle for meaning and is deliberately constructed to transfer information in instrumental fashion. The meaning of each word or symbol in language lies in its signification of the object or relationship to which the word or symbol refers. Claims or statements containing major terms for which there are no empirical referents or statements are metaphysical and, thus, meaningless in scientific terms. Positivists acknowledge that scientific theories may contain propositions that do not have empirical import; however, these propositions are to be considered suspect and may be accepted only so long as the particular statements derived from them can be shown to have empirical import and are confirmed through repeated testings.

This brings us to another aspect of the positivist account of science with particular significance for the social sciences. This is the distinction made by positivism between facts, on the one hand, and values, on the other. Positivists claim that an acceptable scientific (i.e., objective) knowledge claim is a statement that describes some state of affairs as it is, not as it should be. Thus, scientific statements are empirical assertions that can be put to scientific tests. Value claims, which include such valuative terms as "good," "bad," "just," and "unjust" are not empirical assertions that can be shown to be true or false. Thus, as Hempel (1965) noted, "Categorical judgements of value then are not amenable to scientific test and confirmation or disconfirmation; for they do not express assertions but rather standards or norms for conduct" (p. 86). On the positivist account then, classical scientific theories must be descriptive or explanatory in nature but never prescriptive in an evaluative manner.

Although science may not allow or provide valuational judgments, it can provide useful information for making some valuational decisions. Hempel (1965) has noted that

> science can provide factual information required for the resolution of moral issues . . . more specifically, factual information is needed, for example, to ascertain (a) whether a contemplated objective can be attained in a given situation; (b) if it can be attained, by what alternative means and with what probabilities; (c) what side effects and ulterior consequences the choice of a given means may have apart from probably yielding the desired end; (d) whether several proposed ends are jointly realizable, or whether they are incompatible in the sense that the realization of some of them will definitely or probably prevent the realization of others. (p. 92)

Insofar as scientific inquiry produces information useful to decision making, it is instrumental in making relative value

judgments. Nevertheless, on the positivist account it is not possible to derive from scientific theories any categorical value claims.

One of the most ardent expressions of positivism may be found in the "behaviorist" movement in psychology. Fathered by James Watson in the early 20th century, behaviorism posited that only those phenomena operationalized in observable terms were acceptable objects of scientific study. Phenomena incapable of reduction to observable terms and the language of scientific instrumentation were to be considered metaphysical or nonscientific and were to be shunned by the scientific community. Watson maintained that psychology could study humans only in terms of their behavior (i.e., their observable movements) and must reject attempts to inquire into the internal workings of the human mind (i.e., consciousness, motivations). Thus, Watson believed that "only behavior is observable, and only by focusing on this can psychology become objective" (Phillips 1983, p. 6). Perhaps the most succinct summary of the behaviorist program comes from Watson's 1913 paper on behaviorism as quoted by Phillips (1983):

> Psychology as the behaviorist views it is a purely objective experimental branch of natural science. Its theoretical goal is in the prediction and control of behavior. (p. 6)

Behaviorism found expression in other areas of social inquiry as the "behavioralist" movement. The latter was a somewhat weaker rendition of the former and interpreted the operationalist criterion to extend to "verbal behavior" as displayed in responses to tests and survey questionnaires. Thus, behavioralist political scientists could continue to study, for example, the effect that individuals' perceptions of key candidates have on their voting behavior; educators could still objectively determine what students were learning.

An Economic View of the Subject Matter

Since its inception in the early 20th century, vocational education has been justified and rationalized primarily on economic grounds. Moreover, supporters and researchers in the field have tended to interpret vocational education as basically an economic process similar to those found in industrial production. The economic justification for vocational education has been derived from competitive labor market and human capital theories. The economic interpretation has come about from the application of scientific management concepts to the vocational education process. The following sections discuss the background assumptions underlying these theories and concepts.

Human Capital Theory

As Violas (1981) has noted, "The ideas embedded in human capital theory historically have been, and continue to be seen as providing the theoretical underpinnings for public secondary school vocational education" (p. 137). Since human capital theory is an adaptation or extension of competitive labor market theory (CLM), both theories will be discussed briefly.

In the attempt to explain labor market dynamics, CLM posits some assumptions about human behavior. A first and central assumption is that individuals are basically motivated by utilitarian considerations (i.e., the desire to maximize pleasure and minimize pain). Moreover, humans are "rational" beings. Rationality in this context is instrumental in nature. Thus, in using the mini-max approach to pleasure and pain, people will choose those strategies of action that most efficiently obtain the greatest amount of pleasure while reducing as much as possible any resulting pain.

In Western capitalist countries, a person usually needs to be employed or to work in order to obtain such basic necessities as shelter, clothing, and food. Thus, work is perceived as one major method of reducing the pain of deprivation and hardships. Once individuals are motivated to work, they must then decide where to work.

> It is the individual's desire to maximize comparative net advantage which leads to the prediction of this theory that workers will move away from markets paying lower wages to those paying higher wages. (Chamberlaine, Cullen, and Lewis 1980, p. 318)

Although "comparative net advantage" includes the calculation of nonmonetary benefits (e.g., quality of work conditions, fringe benefits), as well as wages, the focus has tended to center on wages because these are measurable and easy to identify.

These assumptions also explain the aggregate behavior of workers because, all things being equal, a given labor force will distribute itself among firms offering higher wages and will attempt to avoid or leave those firms paying comparatively lower wage rates.

These assumptions are used to explain the supply side of labor dynamics. What about the demand side? CLM theory states that individual employers or employing firms are like workers in that they are also motivated by utilitarian considerations. In this context, the maximization of pleasure is synonomous with the maximization of profit, whereas the minimization of pain is the reduction of economic costs. Thus, the attempt to realize net profit guides the behavior of individual employers and firms. A firm, which is also rational in its behavior,

> seeks to combine factors of production in such away as to maximize profits. This requires weighing the marginal cost of a factor of input, such as a unit of labor, against the marginal revenue earned by that same unit, as well as against the marginal costs and revenue of other factors, for example, capital and raw materials. (Ibid., p. 319)

All things being equal, one factor of production may be substitutable over time for another. Thus, machinery may replace labor through automation or robotics. Moreover, the individual firm views workers as being relatively interchangeable and equally efficient. Differences among individual workers are not important or relevant to the firm. The firm will continue to use units of a factor of production just so long as their contribution to marginal value product (i.e., net addition to revenue attributable to the marginal factor unit of production) is greater than their addition to cost.

Numerous criticisms have been directed at the CLM theory; for example, it is too simplistic and its central assumptions rarely hold true in reality. The human capital approach is one attempt to modify CLM theory by bringing it more in line with empirical reality.

Human capital theorists basically accept the CLM framework of demand-supply labor economics. They also accept the fact that individuals are motivated by utilitarian considerations and employ rational, instrumental strategies in pursuing their purposes. They also view labor as one of the factors of production; however, here they begin to diverge from CLM theorists. The divergence arises because human capital theorists think it is a mistake to relegate labor to a status equivalent in importance to that of other factors of production. As Carnevale (1983) forcefully stated, "The human resource is the master economic resource. It acts as the economic catalyst that changes the material and imaginary stuff of our world into usable goods and services" (p. 8). Human resources or capital then has a special role to play in economic growth and development. Becker (1964) further explained this significance:

> Recent years have witnessed intensive concern with the research on investment in human capital. . . . The main motivating factor has probably been a realization that the growth of physical capital, at least as conventionally measured, explains a relatively small part of the growth of income in most countries. The search for better explanations has led to improved measures of physical capital and to an interest in less tangible entities, such as technological change and human capital. (p. 1)

In providing human capital with its special significance, theorists also point out another mistake made by CLM adherents. This

consists of the latter's lack of recognition that labor market success is directly related to differences that characterize individuals within any given labor market. In terms of the human capital approach, the most relevant of these has to do with differences in the skill levels possessed by members of the labor market. This in turn has led to an increase in the attention paid to education and training and their role in the skill acquisition process.

Carnevale illustrated the increased significance attributed to education as a source of skill training:

> In short, productivity turns on human decisions, work attitudes and the contributions of health care, education and training to the quality of the workforce. . . . Growth in on-the-job training, reallocation of labor and the increase in the quality of labor through education, training, and health care have consistently accounted for the lion's share of productivity increases since 1929 and for most of our growth in national income. (1983, p. 41)

From the human capital perspective, expenditures that go to improving the productivity of an individual should be considered "investments." Investments are typically thought of as long-term ventures where both costs and benefits are realized over a number of years. In much the same way, expenditures on human capital that go toward more education and skill training are investments whose "dividends" become significant only over time.

Investments in human capital typically have resulting benefits of three types. The first is the benefit that goes to the individual who, upon incurring the costs of increased education or skill training, will ultimately receive a greater salary or wage rate than would have been received without the additional education or training.

The second type of benefit accrues to the employer or firm who has invested in employees by providing them with on-the-job training or education. In this case, the benefits consist of the difference between the workers' contribution (marginal value product) to the firm and the wage paid to the worker. If the firm does not pass on to workers the full value realized from the additional productivity, the firm will realize a greater net profit from its investment in human capital.

The third type of benefit is that which goes to the society as a whole in terms of greater gross national product (GNP) and increased consumer purchasing power. In short, when investments are made in human capital, "The result will be a market payoff in terms of increased production yielding greater national wealth, corporate profits, and individual wages" (Violas 1981, p. 143).

To summarize, the background assumptions underlying the economic justification of vocational education are as follows: Individuals in the labor market seek to maximize their comparative net advantage. Prior to entry (or reentry) into the labor market, they rationally calculate how to "invest" their resources so as to obtain jobs with greater benefits. The investments are usually in some form of education or skill training. On the demand side, firms seeking to maximize net profit enter the labor market when they view labor as the factor of production that is most likely to yield greater marginal value product (MVP). Moreover, rational calculation indicates that to the extent it has received skill training, labor over time is the greatest single source of continued economic productivity. Thus, firms seek labor educated and trained in the skills needed for efficient production. Together, these assumptions provide a natural locus for the role and function of vocational education--namely, that of training individuals in the skills needed by firms for the efficient production of goods and services.

When vocational education effectively fulfills its role and function, the benefits predicted by human capital theory are expected to result. First, the individual benefits by attaining skilled jobs that he or she would probably not have attained without the training provided by vocational education. The jobs may be entry level; however, entry-level skilled jobs generally provide greater benefits than do entry-level, unskilled jobs (i.e., better wage rates, better fringe benefits) and greater chances for upward job mobility. This is echoed by Evans and Herr (1978):

> One of the primary functions of education and of vocational education in particular is to enable the individual who completes the educational program to enter a career at a higher step, and to progress further along the career ladder than would be possible without the educational background he or she has received. (p. 33)

Second, firms benefit from vocational education because in hiring trained workers, they hire more productive and efficient workers. Although on-the-job training is an investment that can be made by the firm, supporters claim that it can be done more economically by vocational education than by the individual firm.

Finally, society benefits from vocational education because increased employment rates and increased economic productivity result in an increased quality of life for everyone. Using the economic criterion to both explain and justify the need for vocational education has been around since the inception of vocational education in the public school system. In 1914, the President's Commission on National Aid to Vocational Education similar conclusions:

> There is a great and crying need for providing vocational education of this character for every part of the United States--to conserve and develop our resources; to promote a more productive and prosperous agriculture; to prevent the waste of human labor; to supplement apprenticeship; to increase the wage-earning power of our productive workers; to meet the increasing demand for trained workmen; to offset the increased cost of living. Vocational education is therefore needed as a wise business investment for this Nation, because our national prosperity and happiness are at stake and our position in the markets of the world cannot otherwise be maintained. (Panel of Consultants 1964, p. 21)

The scientific management approach. House (1978) noted that the scientific management approach was introduced into American schools between 1910 and 1920. Its introduction was exemplified by the National Education Association's formation in 1911 of the Committee on Tests and Standards of Efficiency of Schools and School Systems. The focus on scientific management was complemented in the 1960s when the Federal Government attempted to introduce the Planning, Programming, and Budgeting System (PPBS) throughout its departments and agencies. The Department of Health, Education and Welfare (HEW) was one of the areas in which the PPBS found strong support. This support was particularly manifested by HEW's creation of the office of Assistant Secretary for Program Evaluation (ASPE), which employed a systems analysis approach to social services. The 1970s also saw a strong push for rationalization of the educative process through legislative activity among the States (Wise 1979). This push was in part an extension of the Federal emphasis on efficient and productive schooling; however, it was also a reflection of the heightened concern for accountability of social programs, including education. Indications were growing that American students were increasingly showing signs of serious deficiency in such basic skills as reading, writing, and mathematics. Rationalization of the schools was perceived as a means to correct this problem.

The scientific management approach was borrowed from the industrial context and reinterpreted for application in the educational setting. It was fathered by Frederick W. Taylor, an efficiency engineer. Taylor saw the need to apply scientific method to the rationalization of the industrial process, especially as it related to human labor. By "rationalization," Taylor meant the maximization of industrial productivity (particularly worker productivity) and the minimization of wasted resources (e.g., human time).

Scientific management literally meant the use of scientific methods by management to "analyze, plan and control the whole manufacturing process in detail" (House 1978, p. 398). Implementation of the scientific management approach could be effected through the following steps:

1. The conduct of scientific time and motion studies to derive precise and specific measurements of the tasks required to perform a particular production process

2. The definition of specific activities and procedures by way of which workers would be told to perform their tasks

3. The development of production standards by which output for each task or phase of the production process could be measured; thereby, identifying efficient and inefficient workers

4. The development of incentive--salary bonuses to encourage efficiency among workers and the punishment, such as firing of inefficient workers whose performance did not improve with incentives or training.

One of the more enthusiastic and vocal advocates of the scientific management approach was Franklin Bobbitt. House (1978) described how Bobbitt viewed the application of scientific management to education.

> It was true, he said, that education was backward and that it had to get new methods from business. It was the job of the supervisors to define the ends of the organization, to find the best methods, and to "enforce the use of these methods on the workers." There were several major principles.
>
> First, definite standards had to be established for each education product, just as in industry. . . . Standards also had to be set for each stage of processing as well as for the final product (the student). Scales of measurement were necessary to know which teachers were inefficient so they could be gotten rid of.
>
> Teacher excuses were no longer acceptable because the measurement scales were scientific and therefore fair. Ideally the product specifications should be set by the public. Also, an inspection system was necessary to make certain the educators didn't cheat.
>
> In education, as in industry, it was necessary to use the most efficient methods. Scientific management would study the best methods, measure the results, select the best methods, and put them into effect. This would take education out of the realm of opinion. Otherwise the supervisor's opinion was no better than that of the teacher's.

> There should be standard qualifications for teachers, and the teachers should be trained to competency specifications by the schools or by the training institutions. Throughout his career, the teacher should be kept up to standard specifications by incentives for professional development.
>
> Management would supply the worker with exact details as to how the work was to be done and with appropriate tools and methods. The worker would be given individual responsibility for each task, and there would be incentives to stimulate output. Management would order and sequence all work processes.
>
> The worker referred to here is of course the teacher. Efficiency, Bobbitt thought, depended on centralization of authority and specific directions by supervisors. The teacher's freedom was less then before, but that was because of "scientific law." (p. 399)

As House (1978) and Wirth (1983) both noted, the systems analysis approach (e.g., PPBS) is very similar to that of scientific management both with regard to method and content. The systems analysis approach views education as a process of social service production. Thus, like scientific management, its aim is to rationalize the educational process in order to maximize efficiency (i.e., increased productivity and reduced waste).

The systems analysis approach posits the following principles:

- Program goals and objectives can be explicitly stated in measurable (i.e., quantifiable) terms.
- The use of scientific methods through controlled experiments can uncover constant and uniform relationships between system "inputs" and system "outputs."
- The findings from such studies can be used to develop formal procedures by way of which schools and teachers can efficiently produce the desired level of program objectives.

Moreover, many advocates of the systems analysis approach believe that a "production function" exists that can identify how much of a given resource (input) will give what increment of a specified output. Lewis (1977) illustrated this point of view:

> What factors create human capital? What is the most efficient method of combining these factors? What is the best technology to use in production at various

> levels and for various outputs? . . . In the economics of education, we attempt to answer these questions by finding the production function for human capital; the technical or physical relationship between inputs and outputs . . . economists wish to know the precise inputs (resources) which enter the production process, the precise relationships between factors within the production process, and the outputs (benefits) which result from these production processes in education. (p. 47)

Competitive labor market and human capital theories provide the background assumptions for an economic justification of vocational education. In addition, the scientific management approach provides an economic interpretation of vocational education as a process.

Thus far, we have discussed the technological framework in terms of its economic and scientific treatments of vocational education as though the two were conceptually distinct from one another. In actuality, the model of scientific knowledge and the metaphor of the industrial production process have merged in the public mind and have become mutually reinforcing. Thus, one finds that the breakup of reality into separate and distinct causes and effects is imitated in the breakup of the social processes into inputs and outputs. The positivist assumption that uniform and causal relations are the "cement of the universe" plays out in the economic treatment as a quest for production functions. As a final example, the role of the scientist as an objective, unbiased observer who accepts only knowledge that is objectively confirmed by rational methods is very similar to that of an "economic man" who chooses action strategies on the grounds that they will most likely produce a desired outcome.

Evaluation Practice under the Technological Framework

The background assumptions and theories previously discussed establish the context within which much of current evaluation in vocational education occurs. They also provide criteria of relevance, which demarcate the appropriate domain for evaluation practice in vocational education. These criteria include the following:

- Human capital theory and the scientific management model provide an economic criterion of rationality. This criterion locates the subject matter of vocational education (and evaluation in vocational education) within the economic domain of industrial production. Thus, the relevant features to look for in accounting for vocational education reality are the extent to which that reality is consistent with economic rationality and rationalization (i.e., the

extent to which vocational education is managed and controlled) so as to produce most efficiently the outcomes for which it is deemed responsible.

- The positivist criterion locates the economic treatment of a subject matter within the domain of scientific inquiry. The relevant explanatory feature emphasized by this criterion is the acquisition or possession of objective knowledge regarding causal relationships in reality.

These criteria are mutually reinforcing insofar as one almost entails the other. Thus, management and control are enhanced by the certainty of knowledge that scientific inquiry can provide. In turn, the acquisition of "objective" knowledge is dependent upon the objective, unbiased rationality displayed by "economic man."* How these criteria and the context they establish relate to evaluation practice in vocational education are discussed next. We suggest that these same criteria demarcate appropriate domains of evaluation practice under the current framework. The demarcation thus serves to restrain response to the following questions:

- What are the major criteria by which vocational education should be evaluated?
- What are the major functions of vocational education evaluation?
- What are the appropriate objects of vocational education evaluation?
- What are the legitimate methods of evaluation inquiry in vocational education?

Evaluative Criteria

Earlier in this manuscript, it is argued that evaluative criteria are inherently related to a framework's background assumptions and criteria of relevance. The latter determine a

---

*Economic man (i.e., rational man) has a subjective desire or motivation. This is to maximize his pleasure or happiness while minimizing his pain or deprivation. However, once you look at means and ends as discrete and separate concepts, you can argue that economic man is totally disinterested with regard to what means are used just so long as they are the most effective (i.e., most productive) in realizing his basic motivation. Thus, the rational calculation of means is a separate act from the valuing of ends.

schedule of human needs, wants, purposes, and so forth and locate those needs within a context for explanatory or evaluative treatment. We have already suggested that the technological framework locates first the need to work and second the need for skills training within an economic context (i.e., the labor market). It follows then that the function and role of vocational education should also be located within the context of the labor market. Within this context, the major criteria of relevance get extended to the primary evaluative criteria for application to vocational education. That is to say, the primary evaluative criteria against which vocational education (and programs thereof) is and should be compared consist of--

- o the degree of productivity demonstrated by vocational education in producing the products for which it is deemed responsible and
- o the degree of efficiency with which vocational education is able to produce these products.

The evaluative criteria of productivity and efficiency operate to define the appropriate realm and scope of evaluation theory and practice under the technological evaluative framework.

Functions of Evaluation under the Technological Framework

One major function of evaluation is to inform decision makers regarding the allocation of scarce resources. As Copa (1977) pointed out, "One of the most formidable problems faced by vocational education decision makers is the difficulty in obtaining adequate resources. Once the financial resources have been acquired, the next problem for a decision maker is the allocation of those resources among the many client groups who need training" (p. 155). Determination of the allocation of resources must rest on other prior determinations. These determinations are made by applying the major evaluative criteria to an evaluation of vocational education and its components (e.g., individual programs). Thus, the primary and more specific functions of evaluation under the technological framework consist of the following:

- o Assessing the productivity or effectiveness of vocational education in producing its products
- o Assessing the efficiency or social rate of return obtained from public expenditures on vocational education

In economic terms, productivity is related to the creation of wealth. Thus, a factor of production is "productive" to the extent that it contributes to the firm's marginal value product

(MVP) and, as a consequence, to its net profit. In vocational education, productivity is also related to the creation of wealth, but the latter is twofold in nature. First, vocational education may be productive in the short term both by making available in the labor market skills that will satisfy the labor needs of industry and commerce and by producing workers who will then be productive within the firm. Within the context, skilled workers may be seen as the "products" of vocational education. "The product of the vocational education industry is a trained person: one who has mastered entry-level skills, upgraded existing skills or retrained for a new occupation" (Hanson and Parker 1977, p. 99). Moreover, within this context, questions of effectiveness become in part questions of productivity (e.g., how many and how much). Thus, assessments of effectiveness determine whether or not vocational education is being productive in producing the types and levels of needed labor market skills that will result in productive workers on the job.

However, there is also a long-term perspective of productivity that is related to the efficiency with which vocational education produces its products. The definition of efficiency is the capacity of "acting or producing effectively with a minimum of waste, expense, or unnecessary effort; exhibiting a high ratio of output to input" (The American Heritage Dictionary 1979, p. 416). In the education evaluation literature, efficiency is assessed through cost-benefit and cost-effectiveness analysis. Hu and Stromsdorfer (1979) explained this perspective:

> A third type of evaluation can be described as cost-benefit or cost-effectiveness and deals with outputs or outcomes, in relation to the inputs or costs involved. . . . This process views the problems of decision making as those of constrained maximization or minimization. Restated, this means that the problem is to choose among feasible alternatives so as to: (1) achieve the most of what is desired with given quantities of resources used, or (2) achieve given quantities of what is desired with the least resources used. (p. 196)

Thus, effectiveness and productivity analysis, along with cost-benefit or cost-effectiveness analysis, serves the primary functions of evaluation under this framework.

Objects of Evaluation under the Technological Framework

The evaluative criteria of productivity and efficiency also operate to establish boundaries for the primary objects of evaluation inquiry under this framework. As was just suggested by Hu ans Stromsdorfer, the primary evaluatio objects of the technological framework are vocational education inputs and outputs or outcomes.

In economic terms, inputs generally refer to the factors of production (i.e., raw materials, capital, and equipment used in the production of goods and services). In much the same way, inputs in vocational education refer to all those things that enter into the production of labor force members trained in occupational skills. Thus, inputs may refer to students' initial cognitive and affective skills, the equipment used in classrooms, teaching and instructional strategies, and so forth. In effectiveness analysis, the inputs that have received the most attention are the goals, objectives, or intended outcomes for vocational education programs. This is simply because the demonstration of effectiveness entails comparing results or consequences with intentions and goals. However, in cost-benefit and cost-effectiveness analysis, inputs are extended to cover all financial and nonfinancial costs that enter into the schooling process.

The second set of objects focused on in evaluation is that of vocational education outcomes. Moreover, emphasis is placed on those outcomes that are economic in nature (i.e., labor market outcomes). Labor market outcomes are taken to be indicators for both the short-term and long-term productivity of vocational education programs. Thus, in looking at the short-term sense of productivity, the outcomes of a number of program completers, the number of student placements in training-related jobs, and the level of employer satisfaction with former vocational education students are investigated. In looking at the long-term productivity or the efficiency of vocational education, outcomes such as job retention by former vocational students, increased or decreased wage rates, and changes in occupational type and level are frequently studied.

The emphasis placed on labor market outcomes as evaluation objects is influenced by the theoretical assumptions and relations that sustain the technological framework and give the criteria of productivity and efficiency their normative relevance. Thus, authors like Darcy (1979) and Grasso (1979) may argue for the importance of noneconomic as well as economic outcomes, but the theoretical priority of the latter combined with the practical consideration of limited evaluation resources indicates that labor market outcomes will remain the most scrutinized objects of evaluation under the current framework.

Methods of Evaluation under the Technological Framework

The appropriate methods of inquiry under the conventional framework are defined by the latter's criterion of what constitutes the "scientific." That is to say, those methods consistent with the positivist model of scientific method are considered to be more credible and thus more legitimate than others in obtaining knowledge regarding the productivity and efficiency of

vocational education programs. Next, we will briefly outline three predominant approaches to vocational education evaluation under the technological framework to indicate how each approach has been defined by reference to both the scientific and economic criteria.

Follow-up studies. Perhaps one of the earliest and most common forms of vocational education evaluation has been the follow-up study. These studies were generally conducted for the informational purposes of State departments and local school districts. The Vocational Education Amendments of 1976 required that local educational agencies throughout the States submit on an annual basis statistics regarding the following:

- The rates of program completers and leavers who secure placement in the occupation for which they were trained
- The level of employer satisfaction with the training and preparation for employment of former vocational education students

Earlier follow-up studies used a variety of methods to respond to these issues ranging from the administration of fairly sophisticated questionnaire surveys to verbal reports based on teachers' personal experience with program graduates and local employers. However, teacher reports and other personalistic methods of follow-up have increasingly become suspect as too unreliable and biased. Thus, the expectation has emerged that more scientific methods be utilized by local school districts and schools for the conduct of student and employer follow-up.

In very general terms, the conduct of a scientific follow-up study includes (at least) the following steps:

1. Construction of a questionnaire that reliably and validly measures the outcomes under consideration
2. Development of a random sample of program completers and leavers and local employers to whom the questionnaire will be administered
3. Standardized conduct of the survey by mail or telephone
4. Analysis of the returned data via statistical or quantitative techniques

In theory at least, the appropriate implementation of these steps will result in unbiased information regarding the vocational education program's level of productivity. Moreover, the study will meet the demands imposed by the positivist criterion of scientific method (i.e., the study will be replicable, its results will be as objective as possible, and the results will be generalizable to the larger population of program completers and leavers and local employers).

Vocational education outcome studies. Like follow-up studies, outcome studies are intended to identify "short-term consequences and longer-term impact resulting from vocational education programs" (Darcy 1979, p. 7). However, unlike most follow-up studies, outcome studies are generally designed on the basis of systems analysis models in which program outcomes are related in some fashion to program inputs. Moreover, outcome studies assume that outcomes can be related to inputs through formulation of a production function equivalent to those used in economic analyses. An example of this is Darcy's (1979) six-factor input-output model where vocational education outcomes are seen as the function of students, context, resources, goals, and processes. Darcy described this model:

> It also resembles a conventional education production function, which relates input factors to educational products or outputs. The five "input" factors . . . which interact to constitute educational programs or "treatments," are recast . . . in the format of a production function. This analogy suggests that over a period of time, and subject to the influence of intervening forces and error factors, it is theoretically possible to observe certain outcomes as having been generated by (i.e., significantly associated with) educational treatments. (p. 9)

Indeed, one of the perceived advantages of outcome studies based on the systems analysis model is that such studies may generate findings that are generalizable across different contexts and time periods. In doing so, they may provide the basis for piecing together cumulative knowledge regarding programmatic cause and effect. One of the major threats to generalizability is the spurious inference of causality. Grasso (1979) discussed this concern:

> In fact, the problem in impact evaluation in vocational education is to produce credible evidence from at least one of several points of view. . . . According to Campbell and Stanley (1963): "Much research in education today conforms to a design in which a single group is studied only once, subsequent to some . . . treatment presumed to cause change . . . such studies have such a total absence of control as to be of almost no scientific value . . . securing scientific evidence involves making at least one comparison."
>
> Whenever this criterion is not satisfied, discussing the findings in terms of impact of the program is not justified. For example, much of the evidence on job placement rates is open to criticism because of failure to consider whether the persons might have obtained the same jobs without any vocational program. (p. 6)

One perceived value of the generalizability of production functions lies in their practical utility. The latter would provide decision makers with a means for predicting the level, quantity, and mix of inputs needed to produce any given level or quantity of desired outcomes. This knowledge could then be used to manage and control the production process more effectively. However, another long-term appeal of production functions is that, if confirmed over time, they may provide a basis for identifying other, broader generalized relationships under which the former can be subsumed. As noted earlier, the possibility of identifying and subsuming lawlike generalizations in a continuing fashion is the cornerstone of the positivist conception of cumulative scientific knowledge.

It should be noted that outcome studies in vocational education have served to identify the extent to which a program can be adjudged as productive, effective, and accountable. An outcome study provides information regarding productivity by simply presenting decision makers with the number of units attained (e.g., a program placed 150 persons in training-related jobs). It provided information regarding the effectiveness of a program by comparing the program's actual performance with its anticipated performance (e.g., a program may be assessed as effective because it attained a 90 percent placement rate). Finally, by providing information regarding productivity and effectiveness, the program indicated whether or not it can account for its actions. However, a more sophisticated indicator of program accountability can be gained through cost-benefit analysis.

Cost-benefit or effectiveness studies. As noted earlier, cost-benefit analysis is usually undertaken to assess the economic efficiency of a program or educational treatment in producing a desired level of outcome(s). In this sense, then, it is not an alternative to the outcome study that determines the effectiveness or productivity of the program or treatment. Rather, it can be viewed as the next step in rationalizing the program development, implementation, and evaluation process. Beyond ex post evaluation, cost-benefit analysis may also be used to estimate anticipated costs and benefits of a program.

Technically speaking, cost-benefit analysis requires that all monetary, and nonmonetary program costs and all anticipated and unanticipated program benefits be identified. Once identified, the costs and benefits are translated into some common measure, usually monetary in nature. Once standardized, "the benefits and costs can then be compared, generally by computing either a benefit-to-cost ratio (benefits divided by costs), the net benefits (benefits minus costs), or some other value (internal rate of return) for summarizing the results of the analysis" (Rossi, Freeman, and Wright 1979, p. 247).

In the case of social intervention programs, the attempt to translate benefits and outcomes into monetary values has proved

difficult. Therefore, the pure cost-benefit model is often revised into a cost-effectiveness model. In the latter, both costs and benefits still need to be translated into some quantifiable measure; however, only the costs need to be estimated in monetary terms. Thus, a cost-effectiveness study of a vocational education program might find that it cost $150 per placement of a student in a training-related occupation. Or, "alternative educational interventions may be analyzed by measuring the educational gains, expressed in test scores, then relating them to program costs" (ibid., p. 248). The point is that in cost- effectiveness evaluation, the substantive outcomes (e.g., job placement, job satisfaction) can be related to the cost for producing them. Programs can then be compared in terms of their efficiency in producing additional units of the outcome.

Without going into detail, we will simply suggest that cost-benefit evaluation methodology is perhaps the most explicit use of scientific and economic criteria. Thus, program costs and benefits are to be operationalized in a standard manner and then quantitatively analyzed. The analysis is expected to be replicable by anyone else and to yield the same results. The analysis is also expected to provide a quantitative representation of the relationship that exists between some mix and level of resources (cost) and the production of some mix and level of outcomes (benefits). The end result of the analysis is to provide decision makers with a measure of program accountability that also specifies in economic terms the efficiency with which the program can produce.

Conclusions about the Technological Treatment of Vocational Education Evaluation

We have attempted in this chapter to provide an explicit reconstruction of the theoretical assumptions and criteria that guide much of current evaluation practice in vocational education. The analysis indicates that, collectively, the economic interpretation and the scientific treatment of vocational education evaluation leads one to certain conclusions. Briefly stated, these include the following:

- Those vocational education programs that are most productive in producing particular economic outcomes (e.g., training-related job placement, employer satisfaction with graduates' job performance, and graduates' ability to retain jobs) are to be viewed as the most effective and most desirable when compared to other programs.

- Those programs that demonstrate the most efficiency in producing the desired economic outcomes are to be given evaluative priority when compared to other programs.

The point is that the scientific application of economic criteria of "goodness" or desirability are to be given theoretical priority in assessing the merits or worth of vocational education programs. Although a number of educators will not be bothered by this, there are others who will be concerned with the primacy of the attention given to economic and scientific criteria. The next chapter summarizes the concerns in this area and discusses their implications for future work in vocational education evaluation.

CHAPTER 3

CRITICISMS OF THE TECHNOLOGICAL TREATMENT OF VOCATIONAL EDUCATION EVALUATION

Evaluation as currently practiced in vocational education has been the target of numerous criticisms. Although criticisms may vary in nature, most imply that current evaluation practice is limited in its contributions to improving the quality of local-level vocational education. Criticisms also implicitly if not explicitly contain suggestions regarding alternative views of evaluation that would strengthen the relation between practice and theory. This chapter discusses the perceived inadequacies in the technological evaluation framework, their relation to practice in vocational education, and the alternative concept of evaluation implied by those criticisms.

Categorical Inadequacies

A major criticism of the technological framework is that adherence to it leads to a fragmentation of any integrated and coherent structure of meaning or knowledge.* It does this by holding to a view of reality that is atomistic, mechanistic, and contingent. Reality, on this view, can be broken down into separate and distinct processes, entities, and domains. The world can be compartmentalized as can our knowledge of it. Indeed, the breakup of moral and social philosophy in the late 19th and early 20th centuries into the separate and distinct disciplines is a reflection of the fragmentation in structures of knowledge encouraged by the technological framework.** This fragmentation has left us with a number of categorical distinctions that were not available to nor employed by pre-Enlightenment thinkers. Thus, distinctions between the sciences and the arts, economic as opposed to political problems, sociological versus anthropological inquiry, and so forth are a "modern" invention associated with the emergence of technological consciousness in the West. It is reflected also in the technological evaluation framework that critics claim has been left with inadequate categories of meaning and inadequate or artificial categorical distinctions for

---

*See, for example, Bowers's (1984) discussion of the modernizing ideology of liberal technocratic educators. In addition, refer to Apple (1983), Taylor (1982), and House (1980).

**For a very provocative (and interpretive) treatment of this thesis, see Cumming (1969).

the explanation and understanding of human and social phenomena. The broader epistemological framework from which these criticisms are derived is discussed in some detail as "interpretivism" in chapter 4. Here, we only summarize in brief the main substance of the criticisms.

Subject vs. Object

In chapter 2, we indicate that the view of reality underlying the positivist account is basically deterministic and contingent in nature. That is to say that occurrences in reality are explainable by reference to some set of regular and uniform relationships between cause(s) and effect(s). Critics emphasize that when this view is extended to the realm of social reality, it completely (and erroneously) does away with any notion of humans as agents who have the capacity to choose in what way they want to live. They argue that in objectifying reality, positivists have also objectified human beings who are perceived to be as much objects of causal determinism as other inanimate or non-human entities. Critics argue that, contrary to this perspective, humans are categorically different from other entities insofar as they are the subjects of their experience and social order. That is, human beings are self-defining subjects who, over time, make choices about the condition of their social environment. Thus, for example, the fact that a given society has a democratic as opposed to an authoritarian political structure must be explained by reference to the political choices that its members have made over time. It is not reducible to explanation by efficient cause and effect in the way that the laws of gravity may be. Thus, the attempt to explain human activity in terms of causal generalizations is perceived to be a categorical mistake that misses the point entirely with regard to the adequate explanation of human and social phenomena.

Meaning vs. Reality

A related criticism aims at overcoming the positivistic distinction between meaning and reality. On the positivist account, meaning is a subjective category imposed on reality by humans. The meaning of a word or symbol lies in its signification of the object or relationship to which the word or symbol refers. Meaning is not, however, a constituent part of the web of relationships and matter that comprise reality. In this view, measurement of our scientific terms becomes essential to ensure that our language is as neutral and objective (i.e., close to experience) as possible. It is only in this way that science can increase the probability that explanations and the meanings contained therein reflect reality as accurately as possible.

However, there is now general agreement among phiolsophers of science and social inquiry that naive or simple empiricism is

an untenable position. Rather, it is recognized that human experience is mediated through the preexisting structures of meaning that individuals possess. Thus, all preexisting knowledge and all our expectations regarding reality are inherently theory laden. As a result, what counts as confirming or disconfirming instances of a scientific hypothesis or theory can not be determined solely by reference to a test result; instead, it will be determined by the interpretations, conventions, and practices that define credibility within the scientific community. In short, knowledge, including scientific knowledge, is inherently related to the meanings that we bestow upon reality.

This argument weakens the force of the distinction between meaning and reality. However, the distinction is completely overridden in some arguments regarding the role of meaning in social reality. These arguments claim that meaning not only mediates reality as experienced by individuals, but it also consitutes most of what we think of as social reality. For example, it is claimed that the reality of modern bureaucracy is constituted by the meanings and interpretations that social members accord certain organizational structures, roles, and practices. Indeed, there is no empirical referent for the concept of bureaucracy outside of or independent of the set of conventions, practices, and values that we interpret (i.e., give meaning to) as a bureaucracy. This is in stark contrast to the assumptions we have that nonsocial reality (e.g., laws of gravity) exists and operates independently of our awareness or experience of it.

Subjectivity vs. Objectivity

Once the distinctions between subject-object and meaning-reality break down, the distinction between subjectivity and objectivity, so basic to the positivist vocabulary, also breaks down. In effect, the elimination of this distinction means that the two edifices upon which the positivist interpretation was built (i.e., formal logic and experience) are not the bottom-line foundations once thought. Instead, formal logic and experience are "creatures" of the meanings and frames of references and conventions through which humans work out their social activities. Does the breakdown of these distinctions mean that all knowledge is arbitrary, that no rational basis exists for preferring one explanation of evaluative conclusion over another? We do not think that this is a necessary conclusion. However, we will discuss these questions more in chapters 4 and 5 when we explore some of the alternative proposals for the foundations of accepted knowledge.

Normative Inadequacy of the Technological Evaluation Framework

Value vs. Fact

The categorical distinctions of subject-object, meaning-reality, and subjectivity-objectivity have had significant implications for the normative stance of the technological framework as well as the normative structure of knowledge generated by that framework. These implications are noted by Smith (1983) in talking about positivism:

> This school of thought claimed that social investigation was a neutral activity in regard to values, and accordingly, social scientists conducting research should (1) eliminate all bias and preconceptions, (2) not be emotionally involved with or have a particular attitude toward the subject, and (3) move beyond common-sense beliefs. (p. 7)

However, with the disintegration of the subjectivity-objectivity distinction, the fact-value distinction built upon it also comes tumbling down. Thus, scientists are not the neutral observers they once were considered to be. They must rely on their stock of social meanings, conventions, and the normative structures underlying these even to be able to make sense of what constitutes the subject matter for investigation. Thus, in chapter 2, we indicate that in employing certain concepts and methods, evaluators have adapted the technological evaluation framework, complete with its normative implications, even when they may not have been aware of it. Interrelated is the argument that non-trivial social explanations (and thus, evaluations) inherently contain some view, however implicity, as to what is to count as desirable with regard to human needs, interests, wants, and so forth. It is nonsensical to claim that a program's having a 10 percent placement rate is a neutral fact devoid of valuative meaning. Indeed, within the technological evaluation framework, the finding only makes sense to us <u>because of</u> its negative value connotations.

Critics argue that maintenance of the categorical distinction between fact and value is not only an error but, more important, encourages the unreflective acceptance of particular normative positions. By attempting to maintain the appearance of "neutral observers," social researchers and evaluators are actually reinforcing or buttressing existing value stances, whatever their respective methodologies (Bredo and Feinberg 1982). We have made the case that within the technological evaluation framework, these normative and political stances imply that productivity and efficiency have priority in determining the worth or desirability of a social program or policy.

Another way in which use of the fact-value distinction has discouraged ethical or valuative discussion is by effectively limiting discussion regarding "validity" to technical discussions about methodological applications. Within this context, validity has to do with measurement instruments and is defined as

> The extent to which differences in scores on it reflect true differences among individuals, groups, or situations in the characteristic which it seeks to measure, or true differences in the same individual, group, or situation from one occasion to another, rather than constant or random errors. (Selltiz et al. 1959, p. 155)

What is omitted from this notion is our commonsensical belief that validity is related to normative appropriateness as well as content relevance and accuracy. In addressing this issue, House (1980) put forth the case of an evaluator evaluating a public program for an external audience:

> In this case it is not enough for the evaluation to be true and credible; it must also be normatively correct. . . . This case is the one that confronts most evaluators. They face triple validity demands that the evaluation be true, credible and right. . . . The task before the evaluator is most formidable because a failure in any one of the three areas invalidates the evaluation. (p. 250)

Once discussion regarding the normative validity of an evaluative claim is closed, there is no longer any basis for open and public debate in this area. This leads to a situation in which particular evaluative claims are taken as given without reflection or debate.

Instrumental vs. Categorical Value Judgments

Earlier, we noted that positivists make a distinction between instrumental and categorical value judgments. Whereas the latter refer to judgments regarding the inherent goodness of a thing, the former relate to empirical knowledge regarding causal relations (e.g., "if A, then B"). Given the value judgment that a policy or program is inherently good and is, thus, desirable, science can provide the necessary technical knowledge to implement or realize the desired outcome. Moreover, science can provide information regarding alternative means for obtaining the end as well as information regarding the most efficient of these means. Thus, on the positivist account, evaluative knowledge is limited to the technical and instrumental. Categorical evaluative questions and arguments are left to be resolved by other (nonscientific) sources.

Several objections have been raised with respect to the instrumental vs. categorical value distinction. One of the major objections is that with regard to social life, the distinction just does not hold up under scrutiny. To indicate why it doesn't, let us review an example used by Hempel (1965):

> Suppose, for example, that careful research had established (1) that restricted upbringing tends to generate resentment and aggression against parents and other persons exercising educational authority, and that this leads to guilt and anxiety and an eventual stunting of the child's initiative and creative potentialities; whereas, (2) permissive upgringing avoids these consequences, makes for happier interpersonal relations, encourages resourcefulness and self-reliance, and enables the child to develop and enjoy his potentialities. (p. 84)

Hempel went on to argue that these findings fall "within the purview of scientific investigation" (p. 84). Does this mean that scientific inquiry has shown that permissive upbringing is a preferred way to raise children and should be viewed as a "good" strategy. Hempel answered in the negative by claiming that all science has shown is that <u>if</u> parents want happy, well-adjusted children, <u>then</u> they should raise them permissively. Or, conversely, <u>if</u> parents desire unhappy, emotionally stunted children, <u>then</u> they should employ restrictive parenting practices. To turn these findings into a prescription for the raising of children, Hempel argued that, you need an absolute, categorical judgment to the effect that one stategy is "good" or inherently desirable whereas the other is not. Since such a judgment would be nonverifiable in empirical terms, it cannot be a claim emanating from scientific inquiry.

Hempel's use of this example is interesting because it works so effectively in a counterargument. Charles Taylor (1974) provides an especially succinct statement of this counterargument. Taylor argues that our everyday language assumes a distinction between the act of judging or evaluating and that of expressing a like or dislike. In social life, there are contexts in which an expression of taste is appropriate, whereas there are other contexts in which a valuative claim is appropriate. Thus, in an appropriate context, a person could make the claim that "chocolate ice cream is bad" with no further grounds for the claim than that he or she does not like chocolate ice cream. However, if a person were to claim that "restrictive upbringing of children is bad," we would expect him or her to adduce reasons underlying the claim. Moreover, we would expect reasons of a certain sort. Suppose the person said that "restrictive upbringing is bad because it results in unhappy, emotionally stunted individuals." Unless there were countervailing considerations, we would find the original claim intelligible in light of the reason or grounds

provided. But suppose the person said that "restrictive upbringing is bad because it increases the number of frowns on peoples' faces." In this case, we would either question the seriousness of the person's claim or expect the person to expand on other, intelligible reasons for thinking that frowns on faces are bad. Intelligence in this context rests on our being able to understand "the judgment as a use of 'good' and 'bad'" (Taylor 1974, p. 55).

Now suppose someone claims that permissive upbringing is desirable because it contributes to human well-being. Taylor (1974) asked, "What can we say here, if asked to give grounds for this affirmation?"

> The answer is that we can say nothing, but also we need say nothing. For that something conduces to human happiness is already an adequte ground for judging it good --adequate, that is, in the absence of countervailing considerations. We come, then to the second point at issue, the claim that to say of something that it fulfills human needs, wants or purposes always constitutes a prima facie reason for calling it "good." (p. 52)

The point here is that in accepting a finding such as "permissive upbringing creates happy, well-adjusted individuals," we must also accept the inherent implication that, all things being equal, permissive upbringing is good. To act in any other way (i.e., denying that permissive upbringing is good) would be to act unintelligibly. The logic of our social language and understanding is such that in matters related to the conditions of our lives, nontrivial empirical claims will have normative implications. Thus, to assert an empirical fact that is nontrivial is also to commend or prescribe. In this sense, then, the distinction between empirical-instrumental and nonempirical-categorical evaluative claims dissolves.

Lack of Evaluative-Normative Dialogue

A second problem with the instrumental vs. categorical value distinction is that one frequently finds the supposedly relative values beginning to assume the significance of categorical values. Thus, in education and vocational education, efficiency and productivity are frequently the evaluative criteria of a successful program or activity. That is, those educational policies, instructional strategies, learning activities, curricula materials, and so forth that enhance program productivity and efficiency and that most efficiently produce the desired level of outcomes, frequently come to be seen as inherently desirable. Thus, House (1978) indicated process-product or input-output thinking "converts questions of educational practice into questions of efficiency. Although efficiency may be relevant as

a goal, it is certainly a distortion of the educational process to take it as the overriding consideration" (p. 400). And yet, this is exactly the crux of the problem. Efficient practices, no matter what their other characteristics may be, cannot be normatively overriden since there are no other evaluative criteria internal to the technological evaluation framework with which to override considerations of efficiency. Surely, critics suggest, there are criteria other than or in addition to productivity and efficiency by way of which evaluations regarding the desirability of particular practices can be interpreted.

Inadequate Treatment of the Subject Matter of Vocational Education

Vocational Education as Training within the Technological Framework

A major inadequacy identified with the economic treatment is the claim that it inaccurately assumes agreement among practitioners and researchers on the appropriate role and function of vocational education. Within the context of human capital theory, vocational education's role and function are perceived to be that of providing society with investment opportunities by training individuals in the specific skills needed by business and industry. The return on this investment is realized through economic outcomes such as greater employment opportunities for graduates, higher salaries for graduates, greater productivity to employers, and greater spending power in the society as a whole. Critics argue however that this interpretation does not adequately reflect the fact that many hold a quite different view of vocational education's major goals and contributions.

Historically, there has been a tension in the way in which vocational education's major objectives have been thought of (Gallinelli 1979). This tension is reflected in the dichotomy that Moss (1983) drew between vocational education as "training" on the one hand and as "education" on the other.

> Webster's Third New International Dictionary, Unabridged (1976) and the Dictionary of Education, 3rd ed (1973) describe education as a process of rearing, bringing up, developing, and fostering growth and expansion of knowledge, and other desirable qualities of mind and character that will be of positive value to the individual in the society in which she or he lives. Training, on the other hand, is seen more as shaping or developing individuals, through drill and discipline, to attain clearly determined, readily demonstrated goals. Education connotes using knowledge and skills to liberate individuals from the bonds of ignorance--to expand her or his options; it is a process done primarily in the

> interest of the person being educated. Training seems to connote using the learning process to mold and control individual behavior toward specific goals/roles that are needed by society; it results in conforming behavior. (p. 20)

Conceptualization of vocational education as a training process is reflected and has been reinforced in Federal legislation by the emphasis placed on "training-related placement rates" and "level of employer satisfaction with former vocational students" as evaluative indicators of vocational education's performance. Nevertheless, there is evidence that large numbers of practitioners and researchers disagree with this formulation and emphasize instead the educational role that vocational education should or does perform. Moreover, the tension between these diverse conceptualizations of vocational education can be found operating within particular programs, within vocational education departments, and within and between local, State, and Federal administrative entities. To the extent that there are differences regarding the substantive content of vocational education, there will also be differences in how people interpret that which constitutes a "good" or strong vocational education program as opposed to a "bad" or weak program.

Noneconomic vs. Economic Dimensions

In evaluative practice, differences in conceptions of role and function also get translated into differences in the priority that certain outcomes receive at the expense of others. Silberman (1983) distinguished between intrinsic or noneconomic outcomes and extrinsic or economic outcomes. The former relates to outcomes that enhance personal development, whereas the latter relates to "employment, income, labor supply, and economic productivity" (p. 6). Operating under the technological evaluation framework, most current evaluation activities focus on extrinsic outcomes. Thus, in their review of nearly 1,500 evaluation studies conducted between 1968 and 1979, Mertens et al. (1980) reported that a significantly higher number of studies focused on economic or labor market outcomes than on noneconomic outcomes. The preponderance of economic outcome studies tends to obscure the fact that many practitioners are attempting to effect noneconomic outcomes. Silberman (1983) made this point:

> Both extrinsic and intrinsic outcomes are important goals for education, but at state and national levels of government (positions that are far away from actual training sites), extrinsic outcomes are more heavily emphasized. . . . As one moves closer to the local training sites where vocational education actually takes place, one finds greater emphasis on intrinsic goals. (p. 7)

The point being made is that when the focus of most evaluation activity is inconsistent with schooling realities, the information generated by those activities is not likely to be relevant or helpful to local practitioners.

Inadequacies in Evaluation Methodology and Practice

Overemphasis on Separate and Discrete Outcomes

In chapter 2, we discussed how the technological evaluation framework has been influenced by the scientific management model of industrial production. This influence has resulted in vocational education's being conceptualized as a production process in which the relation between inputs and outputs can be identified and analyzed.

In practice, however, most evaluation activity is directed toward identifying output or outcome levels. Thus, all of the 50 States plus many local school districts conduct follow-up studies on student placement rates and employers' satisfaction with the performance of vocational education graduates. This focus on output is consistent with the scientific management model's emphasis on performance data as a means of monitoring and controlling the production process. It has also led to a number of criticisms.

Critics suggest that the emphasis on vocational education products or outputs is an oversimplification of what is actually a very complex and multifaceted educational process. Thus, performance data are often considered to be relatively uninformative. Moreover, as Abramson (1979) noted, "It is difficult to draw conclusions about the value of vocational education to the graduates of its programs unless there is some means of understanding the relationship of the program inputs to its outputs" (p. 154). Sirotnik and Oakes (1981) came to a similar conclusion in discussing the use of achievement test scores as the major outcome of interest in educational research: "This testing may signal that something is wrong but the signal is nonspecific as to the genesis of the problem" (p. 165).

The focus on discrete outcomes has also tended to obscure from notice the context within which vocational education is conducted. Traditionally, vocational curricula have been organized around distinct and separate occupational areas (e.g., agriculture, distributive education, and health occupations). Most evaluations are conducted at the level of the particular vocational program in isolation from the broader context of the school. However, a growing body of literature suggests that the

institutional context within which learning occurs has a significant effect on the quality of that learning. As Sirotnik and Oakes (1981) have pointed out:

> One thing is clear: without sufficient understanding of the context within which the process takes place, outcome indices have little or no value, beyond their immediate descriptive signal, for helping direct an agenda for improvement. (p. 165, emphasis added)

Thus, it is felt that in order for evaluations to contribute more meaningfully to local improvement efforts, they must be more explanatory with regard to the institutional dynamics affecting particular outcomes and consequences.

On first glance, it may appear that these criticisms can be defused by simply paying more attention to inputs in evaluations. Indeed, inputs have not been entirely ignored in vocational education. They are a major ingredient in the cost-benefit or cost-effectiveness studies that have been conducted in vocational education. However, inputs considered in such studies are generally restricted to those resources operationalized in monetary or economic terms. Thus, for example, number of hours spent in instruction, the cost of upgraded training equipment, and the cost of resources spent on placement efforts all represent typical inputs in cost-benefit studies. However, the translation of program inputs into other monetary units does little to enhance our understanding of the inner workings of a particular vocational education program. Aside from cost-effectiveness evaluations, there have also been attempts to identify a number of other factors related to programs' successful performance. However, these studies are still frequently perceived to be explanatorily inadequate. This is because most outcome evaluations conceptualize the relationship between inputs as a mathematical function that is describable in quantifiable terms. Thus, for example, one outcome study (Desy, Campbell, and Gardner 1984) reported that once vocational participants graduate from high school, they "are more likely to be active in the labor force-either working or looking for work--than graduates with less vocational concentration" (p. xii). Although this finding may be interesting and even provocative, it does not by itself offer an explanation of why vocational participation and activity in the labor force are associated. Without this explanatory background, it is not clear what local practitioners or stakeholders should make of the finding or how it might be used in improving their vocational education programs.

Overemphasis on the Quantifiable

Not only has the technological framework been critiqued for an overemphasis on program outcomes, but it has also been

critiqued for its limiting evaluation outcomes to those that can be operationalized in quantifiable terms. The emphasis on measurement and quantification derives from the perceived role of causality in relating program inputs to outcomes. As Rossi, Freeman, and Wright (1979) explained, "In the largest context, the problem of discerning the effectiveness of a program is identical with the problem of establishing causality" (p. 162).

As noted in chapter 2, evidence of causality or program effectiveness is sought by use of a quasi-experimental design or some other design that will demonstrate whether or not a causal relationship exists between program inputs and program outputs. Whatever design is used, an essential step in such efforts is to identify the relevant and appropriate measures of cause and effect upon which the evaluation will focus. This is particularly important with respect to outcome measures.

> The starting point for impact evaluation is the identification and explication of one or more outcome measures. . . . An outcome is always a change in the level of measurable variable. (Ibid., pp. 163-164)

Thus, as already noted, the most common outcome measure used in vocational education evaluation is the rate of placement of vocational education students in training related jobs and the level of employer satisfaction with vocational education graduates.

The requirement that evaluation outcomes be measurable in quantifiable terms presents a problem in at least two ways. The first is that local practitioners often need to know about those issues that are not easily reduced to observable or measurable terms.

> It is easier, for example, to measure the number of words that a child spells correctly than to measure that same child's ability to use those words in a meaningful way. . . . It is easier to count the number of minutes a student spends reading books in class than it is to measure what reading means to that child. (Patton 1975, p. 19)

Yet frequently, it is the meaning that instructional activities have for students that is more relevant to local practitioners in their efforts to improve educational quality. Such meanings tend to fall between the cracks in evaluation practice as it is currently conducted.

The second way in which the exclusivity of the measurement principle is problematic is if it begins to actually define vocational education reality. This tendency is related to what Bowers (1982) refered to as the "political nature of language."

> If the world of meaning is shaped in part through images, patterns, organizing principles in our language, then changes in vocabulary will lead to changes in what has meaning, what we are conscious of, and what we honor. The elimination of words can lead to areas of existential silence and political impotence in terms of communicating and negotiation with others over issues affecting lived but inarticulated dimensions of live. (p. 538)

Thus, if vocational educational research and evaluation ignores meanings and experiences other than those that can be translated into observational language, they run the risk of becoming politicized by legislating what may and what may not be entertained in legitimate domains of inquiry and discussion. However, this politicization can be a subtle and implicit process that remains closed off from public or professional scrutiny.

Overemphasis on Scientific Explanation

It might appear that the criticisms of the measurement focus do not intrinsically weaken the position per se. Instead, it looks as though these criticisms are merely stating (1) the need for more sophisticated and refined techniques with which to measure complex behaviors, (2) the need to encompass more independent and dependent variables in research designs, and (3) the need to use statistical analyses that can reflect the interactive and causal relations that exist between independent and dependent variables.

However, we suggest that the criticism runs deeper than this. On the positivist account, the use of measurement not only ensures that our concepts and hypotheses will have empirical import, but it also contributes to our ability to be precise and specific with regard to scientific explanation. In scientific explanation, you explain X by showing it to be a particular instance of the generalization "if Y, then X." The more precisely the relation of Y to Z is described or predicted, the more explanatory your propositions are.

However, explanation in this sense is very much restricted to the technical meaning of logical entailment and empirical import. It has very little to do with our ordinary notions of explanation as a sense-making activity. Explanations based on the logical relations between measurable entities may, in fact, be descriptive of what has occurred. However, they may do very little for us in terms of helping us to understand the meaning of the occurrence. As a television character poignately stated about the death of a close friend, "I know how he died; I simply can't understand why he died." The point is that information that can be integrated into our sense-making activities is what is needed to inform our deliberations regarding practice.

Indeed, critics suggest that the inadequacy of the technological framework in making sense of the performance data is a major reason that decision makers and practitioners pay so little attention to evaluative findings. Thus, as one author noted in discussing attempts to improve the utilization of evaluation findings:

> Researchers have typically responded to utilization problems by attempting to increase the technical sophistication of the quasi-experimental method. This has had the ironic effect of further distancing the world of the researcher from that of the policy maker (Gilsinan and Volpe 1984, p. 181).

These authors go on to suggest that although policymakers appear to be "enamored" with the scientific method, in actuality they rely more heavily on anecdotal (qualitative) data to inform their decisions and planning processes.

Managerial Orientation of the Technological Framework

Another criticism of the technological evaluation framework is that it is oriented toward managerial or supervisory staff rather than to vocational educators and stakeholders. That is, the emphasis on performance indicators results in information being collected that is useful for management issues such as program termination and resource allocation decisions. Although important, these are not the types of issues confronted by practitioners and stakeholders on a day-to-day basis. Thus, an evaluative framework is needed that provides for and allows analysis of information pertinent to the constraints and problems vocational educators must deal with in their classrooms.

Expert Domination of the Techological Framework

Another characteristic of the technological framework that has received much criticism is its tendency to value "expert" knowledge and advice over that of practitioners or local stakeholders. The priority that the expert evaluator receives comes from two sources. The first is the general belief that science as a specialized field with its own vocabulary and methodology generates knowledge that is more "objectively" accurate than knowledge based on individual experience and common sense. The second reason that experts in evaluation are regarded so highly is that "in seeking the most efficient and rational solution the expert appears . . . to be objective and thus above partisan interest" (Bowers 1982, p. 531).

However, critics argue that reliance on the evaluative conclusions and recommendations that experts external to the context

provide can be dangerous. It may be dangerous in the sense that the expert's selection of issues upon which to focus may neither yield a comprehensive nor fair picture of the program. Moreover, an expert-dominated evaluation may not be perceived as empirically and or normatively valid by those with a stake in the evaluation. Thus, any changes or innovations in practices recommended on the basis of the evaluation will run a high risk of not winning the commitment of those who are expected to carry them out.

Agenda for an Alternative Evaluation Framework

Although some of the criticisms regarding the inadequacies in the technological framework are more convincing than others, the sheer number of them as well as the seriousness of the issues involved suggest that, at least, a revised and, at most, an alternative evaluation framework is needed. Taking the criticisms as a starting point, we have developed an agenda of issues that a revised or alternative framework will need to address. The agenda will then serve as a basis for critiquing candidate alternative frameworks in chapters 4 and 5.

Basically, the agenda of issues raised by the criticisms points to the following:

- The need for an evaluative framework that provides an intelligible and meaningful structure of knowledge. By this we mean a structure of knowledge that provides for and stimulates comprehension and understanding. In terms of knowledge, Scriven (1973) stated that

> comprehension or knowledge . . . refers to a psychological state involving knowledge, not of one item, nor of several separate items, but of a field. A field or structure is a set of items related in a systematic way, and knowledge of the field involves knowledge not only of the items but of their relations. Understanding particular items in a field requires knowledge of the relation of the item to other items in the field, i.e., some knowledge of the field. A field is often open-ended in the sense of having potential reference or applicability to an indefinite number of future examples. In this latter case, comprehension involves the capacity to apply to these novel cases the appropriate rule, rubric or concept. (p. 362)

In guiding the generation of such knowledge, an alternative framework should rework or adequately overcome the positivistic distinctions between meaning and reality and between subjectivity and objectivity. Moreover, this reworking should be done in such a fashion that human beings can be regarded as self-formative subjects and not solely as passive objects of contingent forces.

- The need for an evaluative framework that will generate evaluative knowledge. An alternative framework is needed that will recognize the inherent relation between social and human "facts" and values and between objective and subjective in inquiry. One of the major aims underlying this requirement is that implicit normative positions which underly the treatment of a subject matter be made explicit and therefore open to scrutiny and discussion. Thus, an alternative framework is needed that will acknowledge the relation between the normative structure upon which it rests and both its explanatory and evaluative criteria of relevance.

- The need to reformulate the economic interpretation of vocational education. An alternative evaluation framework is needed that extends the economic interpretation of vocational education provided by the technological framework. The reformulation of this interpretation should contain a view of the subject matter that focuses on the "educational" function of vocational education as well as on its "training" function.

  Closely related to this is the need for an alternative framework that lends theoretical and normative priority to the noneconomic as well as economic dimensions of vocational education.

- The need for a framework that provides an interpretation of vocational education as a human activity and not merely a mechanistic process. Within the technological framework, vocational education is treated as (1) some set of separate and discrete inputs that (2) combine and interact in some fashion (i.e., the black box analogy) to produce (3) some set of separate and distinct outcomes. This interpretation has frequently been found to be explanatorily inadequate. Thus, there is a need for a framework that will help us understand the complex nature of vocational education as it relates to the activities (past, present, and future) of stakeholders.

Related to this is the need for a methodology that will allow inquiry into any and all phenomena perceived as relevant to explaining vocational education regardless of whether the phenomena can be operationalized in quantitative terms.

- The need for a framework that is responsive to the interests of groups who are involved in or affected by the activity of vocational education. The framework may allow for the involvement of evaluation or other "experts," but it should attempt to ensure, both theoretically and practically, that such experts do not dominate either the design or the interpretation of evaluation results.

- Finally, an alternative framework is needed that calls for the active participation of stakeholders in defining, designing, and interpreting the evaluation study. Only in this way will findings and recommendations of the evaluation gain the commitment of relevant stakeholder groups.

CHAPTER 4

POSSIBLE ALTERNATIVES FOR AN EVALUATION FRAMEWORK IN VOCATIONAL EDUCATION

In the previous chapter, we laid out a number of the major criticisms of the technological evaluation framework. Furthermore, we suggested that these criticisms imply, if not explicate, an agenda of issues that an alternative evaluation framework would need to address in order to resolve the inadequacies of the technical framework.

In this chapter, we briefly look at major models or approaches that have emerged in the area of evaluation theory. We then conduct a general critique of each of the models with particular attention to how each deals with our agenda for an alternative evaluation framework. The critiques will in turn point out other issues that an alternative evaluation framework will need to address.

In the literature, one can find numerous ways of classifying evaluation models or approaches (refer, for example, to Stufflebeam and Webster [1983]). In this chapter, we will rely with some modification on House's (1980) schema. We chose to do so because one of the distinctions he makes has been central to our discussion thus far. This is the distinction between those evaluation models based on an "objectivist" epistemology and those based on a "subjectivist" epistemology.

Objectivist Evaluation Models

House (1980) used the term "objectivist" in a manner similar to the way in which we have used the term "positivist."

> Evaluation information is considered to be "scientifically objective." This objectivity is achieved by using "objective" instruments like tests or questionnaires. Presumably results produced with these instruments are reproducible. The data are analyzed by quantitative techniques which are also "objective" in the sense that they can be verified by logical inspection regardless of who uses the technique. (p. 6)

House discusses basically four major evaluation models that rely on an objectivist epistemology. These are (1) the systems analysis model, (2) the behavioral objectives model, (3) the decision-making model, and (4) the goal-free evaluation model. The systems analysis model was discussed and critiqued in some detail earlier in this manuscript. Thus, in this section, we will focus on the behavioral objectives, decision-making, and goal-free models.

The Behavioral Objectives Model

The behavioral objectives approach to evaluation is perhaps the most familiar model to educators and vocational education practitioners. It was developed out of the behavioral objectives approach to curriculum design introduced by Tyler (1950) in the 1940s and 1950s. The "Tyler rationale" was closely related to the work of Bobbitt (1918) insofar as it reflected the application of scientific management principles to educational processes.

In brief, Tyler advocated using specific learning objectives as the foci around which curriculum content, structure, and instructional strategies are identified and selected. Influenced by behaviorist psychology, Tyler emphasized that objectives should be defined in terms of the behavioral outcomes students would be expected to demonstrate upon completion of the curriculum (or parts thereof). Explicitly included in Tyler's model of curriculum design is an evaluative component. House (1980) succinctly summarized the aim of the behavioral objectives model of evaluation:

> The evaluation of a program should define its outcomes and its objectives in specific individual behaviors. The task of the evaluator was to determine whether the students were exhibiting these particular behaviors after being in the program. (p. 26)

Thus, the success or failure of a particular curriculum is determined by reference to the extent to which it affects the intended outcomes (i.e., students' acquisition of specific behaviors that have been prespecified).

Evaluations based on the behavioralist objectives model most frequently utilize quantitative measurement of outcomes in the assessment of the latter's "fit" with objectives. Indeed, part of the logic of defining behavioral objectives and outcomes is the enhancement of the "observability" or empirical import of findings. This in turn is assumed to increase the scientific objectivity of the evaluation. Objectivity is also enhanced by the fact that in prespecifying the behavioral objectives for which the curriculum is to be held accountable, the standards and criteria by which the curriculum's success or failure will be assessed are also prespecified. In this way, evaluation is based on scientific principles and on the "facts of the matter," not left to personalistic impressions or subjective bias.

Although Tyler initially conceptualized behavioral objectives evaluation in relation to the design and modification of particular curricula, the model's basic logic and methodology have been extended and applied to the broader context of educational and social program evaluation. Moreover, the behavioral

objectives model is often integrated into a systems analysis approach so that program objectives become interpreted as some of the key inputs comprising the program treatment. The integration of these two models is illustrated by Darcy's (1979) six-factor production model of vocational education and also in Rivlin's (1971) perspectives on systematic evaluation of educational programs.

Critique of the behavioral objectives model of evaluation. In chapter 3, we discuss at length a number of criticisms leveled at outcome-based evaluation. Since the behavioral objectives model is obviously similar in logic and methodology, we will not repeat a full discussion of ways in which it may be criticized. However, we will note those inadequacies as well as the strengths that have been identified with this model.

In terms of strengths, House (1980) noted:

> On the positive side, objectives approaches to evaluation have a great deal of face validity. Should not a social program be held accountable for what it claims to do? These claims, after all, are the basis for public funding of the program. The evaluation has an obvious legitimacy by taking as given enunciated program goals and objectives. (p. 229)

The point is that the behavioral objectives model is more explicit with regard to the standards and criteria to be employed in the evaluation than many of the other evaluation models found in the literature. This at least affords a greater opportunity for public scrutiny of a dialogue regarding the desired objectives of education.

Although the behavioral objectives model is perceived by some to enhance accountability, others claim that it demarcates the evaluation function too narrowly. Indeed, the concern is expressed that the behavioral objectives model relegates evaluation to a purely "technical" activity that is far removed from the activity of valuation. Within the model, evaluation is confined primarily to the technical procedures of collecting data (the nature of which have been prespecified), analyzing the data, and reporting the results of the analysis. Evaluation does not include interpretation of the significance or meaningfulness of the results--there is no need for it to do so since performance standards and criteria are defined quantitatively in the specification of the objectives. Thus, there is no room or need for normative dialogue or discussion in behavioral objectives evaluation. Such questions as, "Are the objectives fair?" or "Are the objectives normatively appropriate?" get omitted from consideration.

This leads to another point. As Scriven (1973, 1980) has repeatedly pointed out, behavioral objectives or "goal-based" evaluation runs the risk of overlooking important evaluative

considerations due to their single-minded pursuit of goal-related information. The narrowness of focus may prevent the evaluator from identifying other unintended outcomes, the consequences of which may serve to undermine or override evaluative conclusions based on the degree of goal attainment achieved by the program. A hypothetical example of such a case would be an evaluation that found a particular curriculum to be successful since students were demonstrating the desired type and level of behaviors upon leaving the program. However, let's assume that the evaluation totally missed the fact that the curriculum content also imparted a number of sexist values and assumptions to students. In a behavioral objectives evaluation, this outcome may well go unnoticed because its manifestations may not be relevant to the intended behavioral outcomes. The point is that the latter outcome (i.e., that of imparting sexist values and attitudes) might serve to weaken or override the judgment that the curriculum is a success <u>if</u> it were ever noticed.

Another area in which the behavioral objectives model can be criticized is in its treatment of the student. Within the model, the concept or understanding of the student as a human being is eliminated. It is replaced by a view that decomposes the individual into separate and distinct behaviors (i.e., a particular curriculum is expected to result in some set of separate and discrete behaviors). Not only does this fragment our understanding of the active learner, it also objectifies the student who is treated as an entity that the program affect causally. The interests, values, and understandings of the student have no role in the behavioral objectives model of evaluation. Thus, how the student perceives the program and what he or she "makes of" involvement in the program are not questions of interest to the evaluator. The latter is only concerned with information pertinent to demonstrating the types and level of behaviors exhibited by the student. This treatment of the student as an object who is of interest only insofar as he or she exhibits particular behaviors is inconsistent with the perspectives of many program stakeholders. Presumably, most stakeholders (e.g., parents and teachers) have a concern for and an interest in the holistic well-being of the individual student.

The student is not the only stakeholder whose perspective is omitted from consideration in the behavioral objectives model. The values, interests, and motivations of other stakeholders are excluded from the scope of evaluative inquiry as well. Their exclusion reflects the managerial orientation inherent in the behavioral objectives model. Its sole concern is to determine the extent to which a program or curriculum has accomplished the tasks with which it has been charged. In addressing this concern, a behavioral objectives evaluation may provide information relevant to the management issues of a "go-no go" sort. However, it will provide little basis for changing or improving program activities or the context within which the program must

operate. Yet, issues of the latter type are more typically of concern to practitioners and other stakeholders.

The Decision-making Evaluation Model

As its label implies, the decision-making model is oriented toward providing information to program planners and administrators for purposes of decision making. Although there have been several variants of the decision-making model (see Provus [1971], and Alkin ard Fitz-gibbon [1975]), we will focus on Stufflebeam's decision-oriented model since it has received widespread attention in educational evaluation.

Stufflebeam (1983) defined evaluation as "the process of delineating, obtaining and providing useful information for judging decision alternatives" (p. 129). Moreover, Stufflebeam stated that

> e CIPP approach is based on the view that the most important purpose of evaluation is not to prove but to improve. It is a move against the view that evaluations should be "witch hunts" or only instruments of accountability. Instead it sees evaluation as a tool by which to help make programs work better for the people they are intended to serve. (p. 118)

The CIPP (context, input, process, product) model assumes that there are basically four institutional or programmatic settings within which decisions must be made. These consist of (1) homeostasis, (2) metamorphism, (3) incrementalism, and (4) neomobilism. However, knowing the decision setting is not enough to ensure that useful information will be provided. Thus, the CIPP model also identifies four decision types that program administrators typically must make. These include (1) planning, (2) structuring, (3) implementing, and (4) recycling decisions. Each of these four decision types requires somewhat different information; thus, the CIPP model incorporates four diverse type of evaluative activities that correspond to the decision types. These are briefly described as follows:

- Context evaluation--the primary aim of which is to "define the institutional context, to identify the target population and assess their needs, to identify opportunities for addressing the needs, to diagnose <u>problems</u> underlying the <u>needs</u>, and to judge whether proposed objectives are sufficiently responsible to the assessed needs" (Stufflebeam 1983, p. 129). Context evaluation usually relates to planning decisions.

- Input evaluation--the primary aim of which is to "identify and assess system capabilities, alternative program strategies, procedural designs for implementing the strategies, budgets schedules" (ibid.). Input evaluation relates to structuring decisions.

- Process evaluation--the primary aim of which is to "identify or predict, in process, defects in the procedural design or its implementation, to provide information for the preprogrammed decisions, and to record and judge procedural events and activities" (ibid.). Process evaluation relates to implementing decisions.

- Product evaluation--the primary aim of which is to "collect descriptions and judgments of outcomes and to relate them to objectives and to context, input and process information and to interpret their worth and merit" (ibid.). Product evaluation relates to recycling decisions.

There are basically three steps involved in the conduct of any of the four evaluation types. These include (1) the delineation of the informational needs or wants of decision makers, (2) the technical activity of collecting the data, and (3) the reporting or dissemination of the information to the intended audience of decision makers. The CIPP model does not prescribe specific data collection methods. Thus, qualitative techniques such as elite interviews and on-site observations may be used as well as quantitative techniques such as quasi-experimental tests or statistical analyses of survey data. However, Stufflebeam noted that all evaluations and components thereof should be expected to meet certain standards. He specifically refered to those standards enunciated in 1980 by the Joint Committee on Standards for Educational Evaluation and published as Standards for Evaluations of Educational Programs, Projects and Materials. The 30 standards are grouped according to the characteristics that acceptable evaluations should possess. These categories of characteristics are utility, feasibility propriety, and accuracy.

According to Stufflebeam (1983), the CIPP model is "geared more to a system view of education" (p. 129). As such, CIPP is not viewed as a one-shot evaluation that has a definite time duration. Rather the CIPP model was developed with the idea of integrating it into the continuous and ongoing decision processes of the institutional system.

Critique of the CIPP model. As House (1980) noted, one of the major strengths and, thus, attractions of the CIPP model is its emphasis on providing useful information that will meet the needs of decision makers. This in turn creates a predisposition in the model toward the generation of context-specific

information or at least information that has relevance for the particular decision-making context and for the particular type of decision. Moreover, the CIPP model quite explicitly allows for the involvement, if not the participation, of stakeholders in each of the four areas of evaluation activity (i.e., in the areas of context, input, process, and product evaluation). Thus, for example, Stufflebeam (1983) advocated the use of stakeholder judgments regarding the program's outcomes in product evaluation as well as more quantitative mesures of outcome attainment.

However, in terms of the concerns noted in chapter 3, the CIPP model is inadequate in several major ways. First, the CIPP model employs a systems analysis approach to the conceptualization of both the evaluation function and that of service delivery. That is, it considers evaluation to be an "input" into the broader service delivery system. As noted in chapter 2, the interpretation of human organizations and activities as "systems" within which separate and discrete inputs and outputs can be identified is consistent with a managerial approach to the rationalization of economic-technological institutions. However, as chapter 3 indicates, many theorists and researchers think that the application of systems analysis categories to the educational or vocational education enterprise is inappropriate. Sirotnik and Oakes (1983) discussed a slightly different issue, but the following statement is still relevant here:

> Technological organizations tightly couple all organizational levels and particularly tightly couple their technical work activities (and outcomes thereof) to their organizational structure. But schools are not technological organizations. In fact, their work activities and products--which we all know to be teaching and learning in at least the four major goal categories of intellectual, personal, social and vocational development--are really at a rather primitive stage of understanding and "technology." (p. 21)

The metaphorical rendering of schools as economic or technological systems tends to give at least implicit priority (both theoretically and normatively) to considerations of efficiency and cost-effectiveness in the "production" of outputs, products, and outcomes (Coomer 1981). This is so despite the fact that Stufflebeam (1983) suggested that the substantive criteria to be used in evaluative determinations may vary according to the issues of relevance to decision makers.

The implicit priority given to technological evaluative criteria is reinforced by the key role that managerial types (i.e., key decision makers) play in a CIPP evaluation. Within the CIPP model, the decision makers enjoy a "first among equals" status vis a vis other stakeholder groups who are involved in or affected by the program's activities. In short, the CIPP model is managerially oriented, as House (1980) indicated:

> Of course, questions arise immediately. Why should the decision-maker, who is usually identifed as the program administrator, be given so much preference? Does this not put the evaluator at the service of top management and make the evaluator the "hired gun" of the program establishment? Does this not make the evaluation potentially unfair and even undemocratic? The answer is that these are potential weaknesses of the decision-making approach. (p. 231)

The point is that managers are typically concerned with evaluation primarily as an accountability tool. Furthermore, in economic-technological organizations, the primary accountability concerns usually relate to questions of productivity, efficiency, and cost-effectiveness. The preoccupation with these concerns obscures or ignores a host of practical and normative issues frequently of much higher priority to stakeholders (e.g., issues of equity and fairness within the organization or in the provision of the organization's services).

Stufflebeam and Webster (1983) claimed that the CIPP model of evaluation "should be distinguished from management information systems and from politically controlled studies because of the emphasis in decision-oriented studies on questions of worth" (p. 33). At one point, Stufflebeam (1983) provided a list of general questions to be addressed by a full implementation of the CIPP model.

- What needs were addressed, how pervasive and important were they, and to what extent were the project's objectives reflective of assessed needs (addressed by context information)?

- What procedural and budgeting plan was adopted to address the needs, what alternatives were considered, why was it chosen over them, and to what extent was it a reasonable, potentially successful, and cost-effective response to the assessed needs (input information)?

- To what extent was the project plan implemented, and how and for what reasons did it have to be modified (process information)?

- What results--positive and negative as well as intended and unintended--were observed, how did the various stakeholders judge the worth and merit of the outcomes, and to what extent were the needs of the target population met (product information)?

"Questions of worth" would appear to be most relevant to responses to the first and last questions. The first question suggests an opportunity for the CIPP evaluator or evaluation to play

a normative role insofar as it speaks of determining the content and priority of client group needs. However, closer inspection of the explanation of the CIPP model indicates that the CIPP model does not in fact determine clients needs. Rather, it identifies what has been ascribed as needs to client groups. Thus, the CIPP model itself does not take a normative stance with regard to the content or schedule of target group needs. In much the same vein, the last question explicates how the CIPP model relates to "questions of worth." It does so by identifying how stakeholder groups adjudge the merit or worth of the program. However, reporting on the evaluative conclusions of various groups or individuals is not the same as making a determination of merit or worth. The former is pretty much an empirical matter. The latter is more complex and involves not only applying evaluative criteria in interpreting what is happening but also adhering to a conceptual or theoretical framework from which the criteria derive. However, Stufflebeam (1983) emphasized that evaluative criteria are to come from the client.

> While the client should at least react to the technical plans, they should exert a major influence in deciding how the findings will be interpreted, e.g., against objectives, against the results of prior needs assessments, based on the evaluator's professional judgment, or through some type of formal group process. (p. 139)

Thus, within the CIPP model, determinations of worth or merit get reduced to the technical activity of applying evaluative criteria that come from sources independent of the evaluation itself.

Based on these considerations, the CIPP model can basically be viewed as an extension of the technological evaluation framework and, thus, not an alternative to it.

Goal-free Evaluation

The fourth model House (1980) associated with an objectivist epistemology is the "goal-free" model, whose primary advocate is Michael Scriven. Two points should be noted at the outset of this discussion. First, the goal-free model has achieved its status as a "major" model both due to Scriven's stature in the field and due to the theoretical issues with which the goal-free model deals and its subsequent influence on the literature of program evaluation. However, it probably has the lowest record of practical application than any of the models discussed in this chapter. Second, as theorists do, Scriven has reworked a number of his major ideas over the years. In reviewing his publications, it is sometimes difficult to identify where some of his themes have been maintained and where older claims have been discarded. As a case in point, Scriven (1977) at one point made as a central component of his argument a distinction between "intrinsic evaluation" and "payoff" evaluation. However, in his

later Evaluation Thesaurus (Scriven 1980) he omitted mention of "intrinsic evaluation," yet has retained the concept of payoff evaluation.*

The goal-free model is very much a reaction against the behavioral objectives or goal-based approach in evaluation.

> Goals are often best seen as inspirational devices--they make poor foundations for analysis. . . . It is also important to note that for evaluators to be aware of the goals of a program is for them to be given a strong perceptual bias in a particular direction which, in conjunction with whatever positive or negative effect they possess for a program, unleashes the possibility of a distorted perception of the results. (Scriven 1983, p.237)

Knowledge of a program's goals, particularly when success of a program is to be judged on how well the program meets these goals, may create bias in the evaluation in several different ways. First, a focus on goals limits the scope of the evaluation to only those indicators of goal achievement. The result of this is that information and knowledge regarding unanticipated results or outcomes will be missed. However, it may frequently be the case that the unintended outcomes are more important to considerations of the program's worth than those originally identified as goals. Thus, Scriven (1973) stated that "it's risky to hear even general descriptions of the intentions because it focuses your attention away from the 'side effects' and tends to make you overlook or down-weight them" (p. 321).

What does Scriven think should replace goals as an evaluation's organizing principle? He (1983) argued that "programs, like products, should be evaluated by matching their effects against the needs of those whom they affect. And this is what the doctrine of 'goal-free' evaluation recommends" (p. 235). Thus, Scriven prescribed that an integral component of evaluation be the conduct of valid and objective needs assessments of consumer or client groups. Results from the needs assessment then play a central role in the determination of a program's merit or worth based on the evaluation findings.

A major concern of Scriven's in his advocacy of the goal-free model was to reduce as much as possible any bias on the part of the evaluator and, conversely, to enhance as much as possible the evaluator's objectivity. However, Scriven's use of the term

---

*We note this only because we will be drawing from assorted works of Scriven and in doing so, may attribute to him or to the goal-free model some claims that may not be consistent with his most recent position.

objectivity was different than that usually found. He claims that the conventional meaning of objectivity refers to "quantitative" objectivity. In other words, conventionally in evaluation, objectivity is equated with the operationalization and measurement procedures used that can be replicated and subjected to public scrutiny. Findings intersubjectively agreed upon as valid are considered to be "objective." "Objectivity in this sense comes to mean that observations are subject to independent verification without reference to the person who produced them" (House 1980, p. 88).

However, this sense of the term objectivity often confuses the issues of validity and reliability. An evaluator may get public consensus regarding the fact that a particular measuring instrument yielded a particular result. Yet, the measuring instrument may or may not be relevant to considerations of program quality. Thus, the instrument may rate highly on reliability but very poorly in determinations of program worth.

Scriven suggested that instead of quantitative objectivity, evaluators need to adopt the notion of "qualitative" objectivity as the standard against which their own inquiry is judged. Qualitative objectivity refers to the extent to which an evaluation accurately portrays the factual state of affairs that exists with regard to the object of evaluation. Moreover, within this context, evaluator objectivity is derived from the evaluator's independence from influences that might distort or bias evaluative judgment. However, on Scriven's account, independence from bias is not equivalent to taking a value-free or value-neutral stance. Indeed, one of the distinguishing characteristics of evaluation is its relation to valuation, as was noted by Scriven (1980):

> The process of determining the merit or worth or value of something; or the product of that process. The special features of evaluation, as a particular kind of investigation (distinguished, e.g., from traditional empirical research in the social sciences), include a characteristic concern with cost, comparisons, needs, ethics, and its own political, ethical, presentational, and cost dimensions; and with the supporting and making of sound value judgments, rather than hypothesis-testing. (p. 47)

Scriven (1977) claimed that there are three types of value judgments. Thus, "some value judgments are essentially assertions about fundamental personal preferences ('matters of taste')" (p. 342). These are factual claims about what a person likes or dislikes. Establishing their accuracy says nothing about what is appropriate or valuable to other persons. It merely establishes the truth of the claim for the particular person. A second type of value claim is "the assessment of the merit or comparative merit of some entity in a clearly defined

context where this amounts to a claim that its performance is as good as or better than another's on clearly identifiable and clearly weighted criterion variables" (ibid.). These too are factual claims. However, Scriven said that establishment of these claims goes beyond merely demonstrating that their claimants believe them--it also makes it "possible to determine whether it is right or wrong for anyone to believe them" (ibid.). Finally, the third type of value judgment concerns the criteria whereby something is "counted as good." This type stands in contrast to the other types, which relate straightforwardly to the "facts of the situation" or, as Scriven (ibid.) put it, to "what is in fact good." Although these may not be factual claims (Scriven suggested that they were instead philosophical claims), they are still subject to rational discourse. Thus, Scriven contended that it was "immature" to avoid them. Rather, "the only proper reaction is to examine the reasons that are put forward for them and see if and how the matter may be rationally discussed" (ibid.).

The fact that Scriven believed value judgments may be derived from rational deliberation (including, in some cases, deliberation regarding the facts of the matter) led House (1980) to say that "his [Scriven's] position is objectivist in that not only the information derived but presumably the standards of evaluation themselves can be objectively determined by the evaluator" (p. 252).

It is pertinent at this point to ask what methdological procedures of inquiry are prescribed in the goal-free model for use in generating objective evaluation knowledge. However, the goal-free model is perhaps more vague on this point than any of the models discussed in this chapter. In his Evaluation Thesaurus, Scriven (1980) noted that "much of the methodology used in evaluation studies is derived from other disciplines--the special nature of evaluation is the way in which it synthesizes these into an appropriate over-all perspective, and brings them to bear on the various kinds of evaluative tasks" (p. 50). Regarding specific procedures, Scriven suggested that "'emergent', 'cascading', or 'rolling'" survey sample designs are examples of evaluation-specific methodologies, as are the "use of parallel teams working independently, calibration of judges, convergence sessions, 'blind' judges, synthesis, bias balancing, etc." (ibid., p. 51). In short, it would appear that Scriven advocated any defensible strategy of inquiry that enhances the objectivity of evaluation findings (i.e., those that protect the evaluator from bias and contribute to obtaining an accurate interpretation of the evaluative case).

Critique of the goal-free model. The goal-free model addresses a major concern that many evaluators and program stakeholders have regarding goal-based evaluation. This concern relates to the "tunnel vision" orientation that many evaluations display in the single-minded pursuit of information regarding the

stated or assumed goals of program or project under study. As has been noted before, the overemphasis on goal-based evaluation is generally managerially oriented at the expense of other stakeholder interests, which may be equally valid. It also runs the risk of ignoring or, at least, glossing over unintended effects of the program that may, in fact, undermine or override evaluative judgments. The goal-free model avoids these dangers by expanding the scope of evaluation to include inquiry into any and all outcomes that may have implications for judgments regarding program worth or merit.

At least initially, it might appear that another strong point of the goal-free model is its emphasis on evaluation as an activity aimed at determining the worth of a program. Moreover, Scriven clearly thinks that determination of worth is something more than simply reporting how particular groups value a program. Yet, closer scrutiny will indicate that although the goal-free model calls for the evaluator to make value judgments, it confines the scope of evaluation to questions of technical worth or merit. This is reflected in Scriven's trichotomization of the types of value claims that may be proposed. The first two value claims discussed by Scriven are said to be factual in nature. Thus, one type relates to claims dealing with personal preferences or taste. These are generally claims that have relevance only for the particular individual who makes them. The second type of value judgment is "an assessment of the merit or comparative merit of some entity" (Scriven 1977, p. 342). It is factual in the sense that it is established by demonstrating that the entity's "performance is as good or better than another's on clearly identifiable and clearly weighted criterion variables" (ibid.). This leaves us with the third type of value claim which, unlike the first two, is not factual in nature but rather philosophical. The third type of claim deals with questions regarding what is to count as good (i.e., it deals with the establishment of evaluation criteria). Presumably, such claims must be settled by philosophical dialogue rather than by appeal to the facts of the matter.

This characterization of value claims is very similar to that provided by Hempel (1965) in his discussion of the relation between scientific inquiry and human values. It will be remembered that Hempel distinguished between relative or instrumental value claims and absolute or categorical value claims. The relative-instrumental claims take this form: Given that X has been determined to be good, policy or entity Y is to be preferred since it performs better than alternatives in creating or increasing the quantity of X. The judgment that X is to count as good is what Hempel called a categorical or absolute value. He believed that only the relative-instrumental claims can be appropriately considered or derived from scientific inquiry. The establishment of categorical or absolute value claims must be undertaken in contexts outside of science (e.g., in the discipline of philosophy). Yet, unlike Hempel, Scriven (1977)

suggested that evaluators should not shy away from philosophical value claims. Indeed, he stated that "the only proper reaction" to such claims "is to examine the reasons that are put forward for them and see if and how the matter may be rationally discussed" (p. 342). However, Scriven undermined the importance of establishing philosophical value claims in evaluation by omitting from the goal-free model any context or location for philosophical examination and dialogue regarding such claims. This omission suggests that Scriven placed a good deal more priority on the establishment of relative value judgments in evaluation than that given to philosophical examination and dialogue regarding such claims. This is reinforced when one considers that in his Evaluation Thesaurus, Scriven (1980) defined evaluation only in terms of an activity aimed at the establishment of factual value claims (i.e., claims regarding the relative merit or worth of a program or entity). We could just let the matter drop at this point. However, Scriven's omission of any context or location for normative debate regarding evaluative criteria presented him with problems in other areas as well.

Scriven's discussion of evaluator objectivity was based on the assumption that the judgment of an "interested" evaluator is likely to be distorted due to bias in favor of the particular interests. However, we have already argued that the generation of knowledge about social reality is of neccessity bounded by the evaluator's interests, values, and theoretical assumptions. In short, the ideal of a "disinterested" evaluator or researcher is, in principle, untenable. We suggest that Scriven is very well aware of this. However, it may be that he is not really so concerned that the evaluator be disinterested as he is that particular interests not be given priority in the evaluation at the expense of other interests. Thus, for example, Scriven made it evident that he believed managerial and program staff's interests frequently predominate in goal-based evaluations.

One indication that Scriven's major concern is a possible predominance of particular interests in an evaluation is the fact that he replaces goals with client group's needs as the foci of the evaluation. Thus, the scope of a goal-free evaluation is to be demarcated by reference to the nature and, presumably, intensity of client groups' needs relevant to program services. In one way, Scriven justified on the basis of needs because they can be objectively determined. That is to say, in using needs as an objective baseline to guide the evaluation, Scriven implied that areas of distortion or bias can more easily be identified. Yet, this is an empiricist argument and Scriven is certainly not a naive empiricist. It makes more sense to interpret his advocating needs as the evaluation standard due to his seeing client needs as more legitimate than the interests of program mangagers and staff. Yet, this still suggests that in giving client needs theoretical and normative priority within an evaluation, an evaluator will approach the subject matter from an "interested" perspective. This brings us back then to the issue of ensuring

that particular interests do not distort or bias the evaluation's validity.

We suggest that this remains a central problem for Scriven, because the goal-free model does not suggest a context within which most or all of the interests relevant to the program can be brought together to examine and critique their perspectives. Indeed, the goal-free model is very much an expert-dominated model insofar as the evaluator is to remove him or herself as far as possible from interaction with program stakeholders. Evaluation activities (e.g., data collection, analysis, and interpretation) are conducted solely by the evaluator. The evaluation's sponsors then receive findings and recommendations based solely on the evaluator's interpretation of the program and its outcomes. However, without a context for comparing alternative "interested" perspectives and for rationally discussing the differences between them, the goal-free model, on its own account, remains vulnerable to the criticism that evaluator interest will bias its results. The point is that the goal-free model needs some context for philosophical dialogue among and between relevant interests in order to address the problems objectively.

Subjectivist Evaluation Models

In contrast to the evaluation models previously discussed are a number of models resting on what House (1980) described as a "subjectivist" epistemology. For reasons we will discuss later, we shall refer to these as "interpretivist" models rather than subjectivist. The central difference between objectivist and subjectivist epistemologies lies in their respective treatments of the nature of reality and of our knowledge of that reality. Whereas the former adheres to a view of an external and independent reality, the knowledge of which is objective by definition, the latter assumes that knowledge is a creation of the human mind and, as such, is inextricably related to the values, meanings, and beliefs that comprise consciousness. Thus, knowledge, from its inception, is inherently subjective in nature.

Interpretivist accounts of social scientific inquiry have received a great deal of attention by scholars in numerous disciplines over the last decade or so. Thus, one finds that Taylor (1982) spoke of "hermeneutical" inquiry, Smith (1983) referred to the "interpretive-idealist" epistemology, Guba and Lincoln (1981) laid out an account of "naturalistic" evaluation, Patton (1980) discussed "qualitative" evaluation methodology, Bredo and Feinberg (1982) critiqued the "interpretivist" approach to educational inquiry, and Bernstein (1978) discussed the "phenomenological alternative."

In many ways, this recent attention may be more appropriately referred to as "renewed" attention. This is because interpretivist accounts of inquiry have existed for quite some time, as

seen, for example, in the writings of Weber and Dilthey in the late 19th century. Indeed, the historical development of interpretivist epistemology may be viewed as a reaction to and against the growing predominance of the technological conceptual framework and, more particularly, its positivistic component. However, it has only been in the last 10 years or so that American scholars have begun to look seriously at interpretivism as an alternative to positivism in school inquiry.

Since interpretivisism represents a major critique of and alternative to positivism, we will devote some time to an analysis of its underlying assumptions. We will then provide a brief description and critique of the evaluation models that rely in one way or another on interpretivist epistemology. These include the case-study or responsive model, the art criticism model, the peer review model, and the quasi-legal model of evaluation.

Interpretivist Account of Social Inquiry

Central to the interpretivist account of inquiry is the view of human beings as interpretive or sense-making beings. This point is explained by Bernstein (1978) in his discussion of Schuetz's work.

> Every wide-awake, grown-up individual approaches his world with a stock of knowledge at hand. . . . He is not meerly a physical being in an objective spatial-temporal world. As a living being who endows his experiences with meaning, he has a position in a world that is meaningful to him. (p. 146)

Conversely, the positing of experiential meaning would appear to entail a view of human beings as interpretive subjects. As Taylor (1982) pointed out:

> Interpretation . . . is an attempt to make clear, to make sense of an object of study. This object must, therefore, be a text, or a text-analogue. . . . In a text or text-analogue . . . we are trying to make explicit the meanings expressed, and this means expressed by or for a subject or subjects. The notion of expression refers us to that of a subject. (pp. 153-154)

The view of human beings as interpretive subjects marks an initial and major departure from positivism because it leads to an opposing conception of the basic category of social inquiry. The interpretivists claim that the general aim of social inquiry is making sense of human activity rather than explaining human behavior in propositional terms. Indeed, non-trivial social research is not interested in the brute facts of, for example, students making marks on paper. Instead, they are interested in the meaningfulness that underlies the observable behavior, for example, of students' test taking or writing skills.

It is due to its intentional nature that action holds meaning for the subject as well as for observers. Moreover, the meaning is derived from the subject's norms, values, and interpretations of his or her situation. It is then informed by the subject's purpose at hand in conducting the activity. Indeed, the reality of an activity being one thing and not another (e.g., test taking rather than doodling) is "defined and identified by the actor's intentions, awareness, and conception of what he is doing" (Winch 1953, p. 87). In short, interpretivists reject the notion that intelligible human activity is caused by external and independent factors. Instead, they argue that activity must be understood by reference to the subject's in-order-to motives and his or her reading of the context within which the activity is to occur. The emphasis on the need for understanding or making sense of human action is an outgrowth of the interpretivist account of concept formation. Briefly stated, the argument runs as follows: All of human knowledge is socially derived and intersubjective in nature. This is because language, the communicator of ideas and concepts, develops and is learned in a social context. The social nature of knowledge points up a major difference between the natural and social sciences. In the natural sciences, the phenomena under investigation has an existence independent of the conceptual systems employed by the investigating scientist. In contrast, the phenomena constituting the content of social inquiry have no existence other than that derived from the concepts and constructs that social members hold and understand. Thus, Schuetz (1953) made the distinction between the "first order concepts," which characterize the study of natural phenomena, and those of the social sciences, which are "constructs of the second degree, namely constructs made by the actors on the social scene whose behavior the scientist observes and tries to explain in accordance with the procedural rules of his science" (p. 303). Thus, the interpretivists hold that the concepts upon which social scientific understanding rests must be consistent with and grounded in the pretheoretical meanings that social relations, activities, and conventions have for the subjects under study.

Their logic of concept formation is carried through to the interpretivist account of explanation in the social sciences. Rather than depending on assumptions of causal generalizations to undergird explanation, the interpretivists usually rely on some notion of "rule-governed behavior" or "typical actions" as the foundations for coherent social scientific understanding. This emphasis is in large part a result of the influence of linguistic philosophers such as Wittengenstein (1971) and Searle (1969). Wittengenstein challenged traditional language learning theories by claiming that the meaning of words and patterns of words is derived from the way they are used in relation to a specific context or within a particular field. Moreover, there are conventional rules regarding the application of particular patterns of words within certain contexts. These include grammatical rules but, more to the point, are the implicit rules of word usage

contained in any public language. Wittengenstein also pointed out that many words are "systematically ambiguous;" that is, the meaning of some words carries the same force independently of the context in which they are spoken. Yet, deeper analysis will show that, although the force of the words remain the same, the meaning of the word may change subtly from one context to another.

Interpretivists often make the analogy between languange learning and human activity. Thus, Louch (1966) noted that "we learn terms in context, and our application of a term is by an analogy to the paradigms which inform our further application of it" (p. 109). Interpretivists extend this notion to apply to our learning of concepts, norms, and, ultimately, rules of conventional behavior. They argue that in attempting to identify and analyze the rules of behavior that are shared by a community's members, social scientists must view those members' actions as expressions of combined symbolic meanings and intentions that the action context holds for subjects. The task of social scientific explanation then becomes that of interpreting, making intelligible, or making sense of the action or activity in terms meaningful to both the subject and the observer.

Beyond the general nature of explanation, the interpretivist position has siginficant implications for the kinds of propositions employed in social scientific interpretations. Winch (1958) indicated this in his discussion of the method social scientists use to identify regularities upon which to base generalizations:

> The concepts and criteria according to which the sociologist judges that, in two situations, the same thing has happened, or the same action performed, must be understood in relation to the rules governing sociological investigation . . . whereas in the case of the natural scientist we have to deal with only one set of rules, namely those governing the scientist's investigation itself, here what the sociologist is studying, as well as his study of it, is a human activity and is therefore carried on according to rules. And it is these rules, rather than those which goven the sociologist's investigation, which specify what is to count as "doing the same kind of thing" in relation to that kind of activity. (p. 87)

The point is that social scientists can make generalizations about human activity, but these generalizations must be cast in terms of the conventional meanings and rules that guide the behavior in question. Moreover, unlike the causal generalizations sought by positivist inquiry, the application of rule-informed generalizations is only valid for the society or community within which the rules have been generated and, thus, taken for granted by members.

Smith (1983) summarized three key areas in which the interpretivist account (referred to by Smith as the interpretive-idealist or I-I position) is in sharp disagreement with the technological-positivist account (referred to by Smith as the quantitative-realist or Q-R position). The first of these has to do with the relationship of the investigator to the investigated. On the Q-R account, the objects of investigation are separate and distinct from the scientist. Thus, "the objects of study will exist prior to an investigator's interest in them, and the investigator's task will be to describe these objects and the relationships between and among them" (p. 43). In this view, is a separation between the objects of study (i.e., human behavior) and the subject of the study activity (i.e., the investigator). On the I-I account, this subject-object dualism breaks down because the investigator is involved in the same activity as that of the human beings under study (i.e., he or she is involved in the active interpretation of human activity just as the human subjects are actively interpreting the situation in order to act). Smith (1983) noted:

> It is impossible, given that reality is mind-dependent, not to have the investigator and the process of investigation influence what is investigated. Whereas in the former case, the investigation responds to an external, independent referent, in the latter the process is internal in that the investigator and the investigation actively participate in shaping reality. (p. 43)

In short, on the interpretivist or I-I account, the investigator and the investigated are human subjects who must actively participate in a common meanings context in order to comprehend and be comprehended.

A second area in which interpretivism conflicts with the technological-positivist framework is in its claims regarding the relation between facts and values. In chapter 2, we discussed the positivist characterization of objective knowledge.

> From the Q-R perspective, the objectivity of an observation must have its referent in the world of facts or reality that exists independent of the knower. An investigation can then be judged objective of the process, and the results can be seen as undistorted or unbiased by the particular interests, dispositions, or situation of the inquirer. (Ibid., p. 45)

However, the interpretivists reject the fact-value distinction by rejecting the positivist distinction between statements and sets of statements having empirical import (i.e., scientific statements) and metaphysical statements having no empirical referents, which are thus nonverifiable or nonfalsifiable. Indeed, interpretivists claim that all of what is taken to be knowledge is influenced by "metaphysical" or subjective interests, assumptions, predispositions, and so forth. These are similar to the

prior interpretive schemas of the nonscientist who approaches life each day with a stock of experiential and "common" knowledge. In short, there is not an external vantage point or "God's-eye" view from which the scientist can directly, without mediation, apprehend the objects of study. This was summarized by Smith (1983):

> This is what Dilthey meant when he said that one cannot stand outside of history or outside of the events of life and then presume to say what they signify. Investigations of the social world are socially and historically bounded; values and interests shape how reality is discussed, studied, and thereby shaped. (p. 46)

Moreover, interpetivists argue that human activity is only meaningful and can only be understood as meaningful when it is comprehended in its normative as well as its empirical dimensions. As Taylor (1974) has argued, our explanations of and, thus, understandings about human activity will entail implications for valuation since they will ultimately and inextricably be cast in terms of some conceptions of human well-being.

Finally, Smith (1983) made a statement concerning the interpretivist perspective:

> Social inquiry is only meaningful because it does involve values. The study of human beings is the study of moral actors--that is, people acting on the basis of their own values and dispositions. To adopt a detached attitude or a God's Eye view (assuming this is even possible) would only be a failure to understand what social inquiry is all about. (p. 74)

The Case-Study Model*

The case-study model has perhaps been the evaluation model that most explicitly relies on the interpretivist epistemology. Stake (1978) gave an example of this in his statement:

---

*Actually, the case-study approach to inquiry is utilized in a number of diverse fields of inquiry (e.g., history, medicine, and comparative politics). In its broadest sense, it simply refers to the study of a particular case rather than some aggregate of cases. Used in this way, the term case study is not intrinsically related to any particular epistemological position. However, within the context of educational research and evaluation, the case study has most frequently been used to refer to efforts that draw off an interpretivist epistemology and that have the characteristics described in this section.

> I claim that case studies will often be the preferred method of research because they may be epistemologically in harmony with the reader's experience and thus to that person a natural basis for generalization. (p. 5)

House (1980) noted that "the aim of the approach is to improve the understanding of the reader or audience of the evaluation, primarily by showing them how others perceive the program being evaluated" (p. 39). He went on to make the point that when the "aim of inquiry is understanding rather than explanation and propositional knowledge, the case study is often superior to other modes of inquiry" (ibid., p. 40).

Advocates of the case-study approach argue that clients or sponsors of evaluations usually want and need information regarding a particular entity (i.e., a particular program, a particular curriculum, a particular institution). Based on interpretivist assumptions regarding the inherent relation between meaning, intentions, and human action, the case study is more likely to provide the detailed, holistic understanding of the situation than are more conventional modes of inquiry.

Although the case study is generally thought of as a "qualitative" approach, it need not be confined only to qualitative or interpretive procedures of inquiry. However, within the context of the case study, the use of quantitative data is likely to be limited to that of "cueing" the evaluator to implicit problems or issues that may be explored more thoroughly through in-depth interviews, on-site observations, and so on. Thus, for example, in a case-study evaluation of a vocational education program, the finding that the program consistently displays low job placement data may be used as a beginning point for exploring the nature of the program's relations with local employers as well as student reactions to the program's job placement services. The point is that the case study rejects the notion that quantitative measurement is the portal to reality.

Moreover, the case-study approach is usually more interested in reporting how stakeholders perceive and evaluate the case than it is in attempting to define the true state of affairs in the case. Indeed, it has sometimes been claimed that the search for an account of a single reality is a pointless and, in principle, impossible task. Guba and Lincoln (1981) illustrated this point with their claim concerning naturalistic inquirers:

> They focus upon the multiple realities that, like the layers of an onion, nest within or complement one another. Each layer provides a different perspective of reality, and none can be considered more "true" than any other. Phenomena do not converge into a single form, a single "truth," but diverge into many forms, multiple "truths." (p. 57)

In rejecting the notion of a single independent reality, case study methodologists often state that their purpose as evaluators is not that of social scientists. The latter aim to provide explanations of the phenomena under study, whereas the former seek to promote empathy with and comprehension of the intentions and activities of the actors involved in the case. Stake (1978) elaborated on this view:

> Polanyi distinguished between propositional knowledge and tacit knowledge. Propsoitional knowledge--the knowledge of both reason and gossip--was seen to be composed of all interpersonally sharable statements, most of which for most people are observations of objects and events. Tacit knowledge may also dwell on objects and events, but it is knowledge gained from experience with them, experience with propsositions about them, and rumination.
>
> Tacit knowledge is all that is remembered somehow, minus that which is remembered in the form of words, symbols, or other rhetorical forms. . . . Tacit knowledge includes a multitude of unexpressible associations which give rise to new meanings, new ideas, and new applications of the old. . . . Explanation belongs more to propositional knowledge, understanding more to tacit. (pp. 5-6, emphasis added)

In attempting to provide the evaluation audiences or sponsors with increased understanding, the case-study evaluator must first be sure that he or she is obtaining an appropriate reading of what stakeholders think and feel. This means that the "design" of a case-study evaluation will be quite different from that of the more scientifically oriented study. Stake (1983) suggested that the latter adopt a "preordinate" approach (i.e., variables of interest are determined, data collection techniques including measurement strategies and data analysis techniques are all specified prior to the evaluation). It is anticipated that no, or only minor, modification of the design will be necessary as the evaluation proceeds. In contrast, the "design" of the case-study evaluation will typically emerge or evolve as a result of interactions with program staff and stakeholders. Thus, Guba and Lincoln (1981) stated, "The approach requires constant interaction between evaluator and relevant stakeholding audiences. The evaluation process moves through several iterations; it makes credibility checks possible at each stage of and invites negotiation on points of differences" (p. 381). Issues of relevance to the evaluation are expected to unfold as the layers of experience (i.e., the multiple realities) within which program activities are conducted make themselves known to the evaluator. "As the program moves in unique and unexpected ways, the evaluation efforts should be adapted to them, drawing from stability and prior experience where possible, stretching to new issues and challenges as needed" (Stake 1983, p. 303).

In attempting to enhance our tacit knowledge about and understanding of the case in question, the presentation of the evaluator's findings will frequently differ greatly from the conventional technical report given to sponsors. In conventional or preordinate evaluation reports, findings will typically include a listing of conclusions about "the facts of the case" based on the evaluator's analysis of the data. These conclusions may be accompanied by or include the evaluator's recommendations regarding specific or general strategies for improving activities or for resolving problems. In contrast, a case-study report will often focus on providing the reader with insight into how key actors or stakeholders view and assess the case. This is indicated by Stake (1978):

> Most case studies feature: descriptions that are complex, holisitc and involving a myriad of not highly isolated variables; data that are likely to be gathered at least partly by personalistic observation; and a writing style that is informal, perhaps narrative, possibly with verbatim quotation, illustration, and even illusion and metaphor. Comparisons are implicit rather than explicit. (p. 7)

The point is that unlike conventional evaluation reports representing only the evaluator's or evaluation team's conclusions, the case-study report will reflect and attempt to illuminate the multiple realities that reside within and among actors on the scene. Weiss's (1983) remarks about stakeholder evaluation also apply to the case-study model:

> It realizes--and legitimizes--the diversity of interests at play in the program world. . . . It takes evaluation down from the pedestal and places in in the midst of the fray. It aims to make evaluation a conveyer of information, not a deliverer of truth; an aid, not a judge. (p. 11)

The nature of the evaluative judgments and recommendations reported in a case study will often differ significantly from that found in more technical reports. In the latter, evaluative assessments will usually issue from the evaluator's interpretation of the data and some set of predetermined evaluative standards. In other words, the evaluator will be sole judge and jury of the program or evaluative object. In contrast, to the extent that a case-study report contains evaluative judgments and recommendations, the latter will reflect and derive from negotiations between and among the evaluator(s) and the various stakeholder groups.

The case-study approach to evaluation is generally perceived to be "responsive" in nature on at least two counts. First, it is reponsive in the sense that it aims at satisfying the informational needs of both the evaluation's sponsors and the relevant

stakeholder groups. Second, it is responsive in the sense that it ensures that stakeholder perspectives, values, and interests are integrated as an integral part of the evaluation. While discussing the "stakeholder" approach to evaluation, Weiss (1983) also described some of the responsive characteristics that have been attributed to the case study model.

> By empowering an array of concerned groups to play an active part in the evaluation, the stakeholder approach makes fairness a central tenet. Since it gives high priority to their definitions of information requirements and to timely feedback of the information that they want, it democratizes the evaluation process.
>
> The stakeholder approach should also improve the relevance of information, because it gives stakeholders a strong voice in deciding what information shall be collected. In the process, and by their very diversity, stakeholders will help to avoid the narrowness of measurement that occasioned many complaints. (p. 8)

Critique of the case-study model. The case study approach to evaluation has gained some popularity in the field within the last decade or so. Its rise in popularity has paralleled the emergence of some disenchantment with evaluative efforts in general. It has been claimed that evaluation has fallen into disrepute due to the fact that many evaluation studies have been too "narrow," "unrealistic," "irrelevant," "unfair," and "unused" (Weiss 1983). At the core of this criticism has been the general opinion that "evaluators too often select for attention the issues that are easy to study with available social-scientific tools. Evaluators . . . choose their issues and variables for reasons that have little to do with the needs of people who make decisions about the program" (ibid., p. 4; refer also to Gilsinan and Volpe [1984], and Patton [1978]).

Its adherents believe that the case-study approach in its reponsive mode poses an effective counter to such criticisms. Clearly, the case study has the potential for providing useful and relevant information since it draws its criteria of "useful" and "relevant" from both evaluation users and stakeholder groups. Moreover, case-study evaluations are pluralistic in that they give equal consideration to sponsors' (managers and administrators) and to stakeholders' informational needs. Thus, unlike the decision-making model, the case-study evaluation does not easily turn into an adjunct of management. Responsive case-study evaluations are considered to be fair because they attempt to give each and every stakeholder group an "equal hearing" with regard to perceptions of the program, valuations of the program, and the formulation of recommendations for action. Finally, the case-study approach includes as a central methodological task the rendering of information regarding the program into holistic understanding of the program, its context, and the activities comprising it.

Although gaining in popularity, the case study has also been the target of much criticism. Much of this criticism has been of the kind one would expect from a predominantly "scientific" community (i.e., case-study methodology is too subjective and nonverifiable). However, one area of significant concern to sympathizers as well as opponents of the case study is the latter's apparent discarding of any basis for dealing with questions of preferred knowledge. That is to say, some case-study theorists have provided a treatment of interpretivisit epistemology that appears to accept a relativist position. In rejecting the notion of a single, objective reality, they talk instead of "multiple realities," none of which can be considered more true than another. This is, in part, the basis for including in case study reports a characterization, portrayal, or discussion of the multiple perspectives of various stakeholder groups. Since these perspectives reflect the multiple realities experienced by program stakeholders, no single one can be considered more valid or accurate than another; hence, the strategy is to present them all. Yet, this account of multiple realities and perspectives tends to confuse the interrelated but distinguishable notions of the reality within which programs exist and descriptions of or claims about that reality. As Scriven (1983) pointed out:

> The concept of relativism is self-refuting; if everything is relative, then the assertion that everything is relative cannot itself be known to be true. So, although we may reject the existence of a single correct description, we should not abandon the idea that there is an objective reality, though it may be a very rich one that cannot be exhaustively described. It may even be one which can only be described in a nonmisleading way by giving descriptions which are relative to each audience; we may concede all this, and yet insist that in many cases there is such a thing as a correct--though not unique--description . . . by contrast with a number of incorrect descriptions. (p. 239)

Actually, the fact that a program may be understood and described in different terms is unproblematical until we have a case where the alternative descriptions contradict one another. For example, one stakeholder claims that a program is worthless because its services are of such poor quality and should not be considered services at all. Another stakeholder makes the opposite claim that the program is highly valuable because it provides a number of client groups many services that fulfill real needs. Ultimately, we can show or come to some agreement that one or the other of these descriptions is totally or in part erroneous. Indeed, the very concept of evaluation incorporates a notion of coming to some determination. Generally, making determinations regarding knowledge claims means acceptance based on considerations of accuracy as well as appropriateness and other criteria. Phillips (1983) summed up the implications of a relativist position.

> After all, not everyone who claims to know actually does so; I may claim my mother had purple hair, or that I know Mozart wrote Oliver Twist, or that I know the meaning of Tchaikovsky's Fourth Symphony, but it is conceivable that I am mistaken, so there has to be some ways of validating or refuting these claims. To deny this is to do away with the concept of a mistake--a heavy price to pay in support of antipositivism. (pp. 11-12)

The point is that in rejecting empiricism as a basis for determining the truth of knowledge claims, one does not have to reject the notion that grounds exist for preferring one claim or set of claims over others. Moreover, as indicated in chapter 5, the acknowledgement that social reality is constituted by intersubjectively shared meanings and understandings does not lead to the conclusion that any single account of that reality must be accepted as a valid or true interpretation.

Other case-study theorists do not go so far as to say that the multiple perspectives of stakeholders reflect multiple realities. Rather, they take the position that all viewpoints regarding the program should be represented without the evaluator's making judgments regarding their respective validity. Thus, the "disinterested" evaluator, as House (1980) noted "should not favor any particular interests in the evaluation, but should treat all the same. In this view, the evaluator neutrally represents and portrays, leaving the weighing and balancing of the various interests to the judgment of the reader. The evaluator makes none" (p. 245). However, as we have already argued, the evaluator can never be totally "disinterested" in the generation of knowledge. Moreover, House (1980) pointed out, "People being evauated do not want a neutral evaluator, one who is unconcerned about the issues. . . . Being disinterested does not give one the right to participate in a decision that determines someone's fate to a considerable degree" (p. 92).

In adopting the interpretivist epistemological base, case-study advocates become vulnerable to another major criticism. That is, in providing an alternative to positivism for the explanation of human behavior, interpretivists have gone too far in eliminating entirely the role of causes. Just as positivist accounts make a conceptual error in omitting reasons altogether from their explanations of behavior, the interpretivist case-study theorists also make a mistake in thinking human conduct must be understood only in terms of conscious motivations, intentions, or reasons. MacIntyre (1982) forcefully presented this argument in his critique of Durkheim and Winch.

> Clearly if the citing of reasons by an agent, with the concomitant appeal to rules, is not necessarily the citing of those reasons which are causally effective, a distinction may be made between those rules which

> agents in a given society sincerely profess to follow and to which their actions may in fact conform, but which do not in fact direct their actions, and those rules which, whether they profess to follow them or not, do in fact guide their acts by providing them with reasons and motives for acting in one way or another. (p. 297)

As an example, MacIntyre focused on role concepts since they are often used by interpretivist sociology in their explanations. MacIntyre argued that there are some roles in which considerations of causality need not arise. Now, actors in these roles do conform to certain norms so that their behavior may be described as rule governed. However, the rules in this case will be similar to those governing the role of a chess player. "We are not disposed to say that the rules of chess or the norms which define the role of a headwaiter constrain the individual who conforms to them" (ibid., p. 299). However, there are other roles in which the network of rules and norms do in fact constrain and affect the behavior of individuals and the options for conduct of those individuals. MacIntyre (1982) cited an example of this from Goffman's (1961) research on role arrangements in mental hospitals, monasteries, and the armed services.

> Goffman concludes that the behavior of patients is determined to a considerable degree by institutional arrangements which provide a severely limited set of possible roles both for patients and for the doctors and orderlies with whom they have to deal. Thus, the behavior of individual patients of a given type might be explaned as the effect of the role-arrangements open to a person of this type. (p. 299)

The point is that the actions of mental patients could be intelligibly explained by reference to the effects on them of the institutional role arrangements; this was true despite or, perhaps, in lieu of any of the reasons that patients provided for their actions. Thus, MacIntyre concluded that Durkheim and Winch both commit conceptual mistakes by not recognizing that both causes and reasons may be necessary for adequately explaining human conduct.

Judicial Evaluation Methodology

As House (1980) noted, quasi-legal procedures and "blue-ribbon panels" have long been used as a means of investigating and evaluating public issues. The Warren Commission, the Kerner Commission on Civil Disorders, the National Commission on Excellence in Education, and the National Commission on Secondary Vocational Education are just a few examples. It is therefore not surprising that variants of such quasi-legal approaches have made their way into the policy and educational evaluation

literature. In education, Wolf is perhaps the most vocal advocate for judicial evaluation efforts. Thus, we will focus on his work to analyze and critique judicial evaluation methods.

Wolf (1979) claimed that conventional policy evaluation efforts stand in need of methodological reform. This need is most apparent in two areas. The first area is the need to allow groups likely to be involved in or affected by the policy issues in question to participate in the evaluation. The second area is the need for an evaluation methodology that will inform policy formulation and not focus solely on policy outcomes post facto. According to Wolf (1979), the judicial evaluation method is an attractive alternative to conventional evaluation methods insofar as it addresses these two needs. Thus, he claimed that the judicial evaluation method represented:

> the articulation of an evaluation method that allows for the systematic and legitimate involvement of parents, teachers, citizen groups, taxpayers, school board members, students and administrators. (p. 20)

Secondly, Wolf (1979) said the judicial method "is directed toward clarification and understanding of the complexity involved in educational programs and the subsequent offering of recommendations that can inform responsive action" (p. 22).

It is a mistake, according to Wolf, to view the judicial evaluation method as a strict or literal application of trial-by-jury procedures to the evaluative enterprise. It is also a mistake to view it as an adversarial debate in which opposing viewpoints confront each other in hopes of winning or scoring points. Rather, the judicial method should be viewed as an application (where appropriate) of the metaphors of law where such application may be useful and appropriate. Instead of winning a debate, the point of the judicial method is to "establish systematic procedures for inquiry and set forth criteria for classifying, evaluating, and presenting evidence in a clear, cogent, and reasonable manner" (Wolf 1979, p. 21). In doing so, the judicial method will be educational for all parties involved since it will provide "reasoned and reasonable discussion" of the diverse perspectives relating to the policy in question.

In operational terms, the judicial evaluation method consists of four stages of activities. The first stage is that of "issue generation." This is an exploratory stage to identify a broad range of issues related to the policy or decision in question. Wolf (1979) defined the word "issue" in this context as "a statement or propósition upon which ordinary persons could disagree and thus establish arguments around points of disagreement" (p. 26). The issues identified are expected to reflect the concerns of various stakeholders relevant to the policy in question.

The second stage is that of issue selection. Here, the pool of issues identified in stage one is reduced to those with the

most relevance to or saliency for stakeholder groups. This reduction of issues to a manageable number is necessary to delimit the scope of the inquiry and, thus, to define the dimensions of the policy most in need of deliberation. The issues selected in this stage serve as the focal points for building and presenting arguments. This leads us to stage three.

The third stage consists of the development of the case and the preparation of arguments to be presented in stage four. Case building and argument preparation include activities aimed at gathering and documenting evidence. These in turn include any and all procedures for data collection and analysis (e.g., surveys, in-depth interviews, on-site observations). In this stage, potential witnesses for the case are interviewed and their "testimony" integrated with other evidence. Eventually, the evidence and the conclusions and inferences derived from it are constructed in the form of an argument for interpreting and acting on a policy in some particular manner.

Stage four is establishing and conducting the clarification forum. The latter includes the creation of a panel of representatives of stakeholder groups. The function of the panel is to hear the arguments regarding the policy and relevant issues. The presentation of the arguments is made public in this way. The forum itself has a format that "approximates that of a court of law" (Wolf 1979, p. 22).

> Case presenters make their cases through witnesses selected to represent their views relative to a given issue. Direct, cross, re-direct, and re-cross examination of all the witnesses are engaged in. And, as in a court of law, opening and closing arguments are presented. (Ibid.)

The panel may ask questions of witnesses in order to clarify points. After the closing arguments are given, the panel members deliberate among themselves and then provide their recommendations regarding the policy in writing.

Although any number of appropriate inquiry procedures may be used in generating evidence, the judicial method is most explicitly dependent on those methods characteristic of naturalistic or case-study evaluation.

> In contrast to more "scientific" methodologies, which generally exclude human testimony, and judgment in the spirit of seeking objectivity, the "legal" model places a premium on these forms of evidence. . . . The ultimate evidence, then, which guides deliberation and judgment includes not only the "facts," but a wide variety of perceptions, opinions, biases, and speculations, all within a context of values and beliefs. (Wolf 1979, p. 21)

Although human testimony may be the "ultimate" evidence, the judicial evaluation method does not rule out other data collection methods, either quantitative or qualitative. Thus, for example, findings from correlational analyses may be used with expert testimony in presenting a case to the clarification panel.

Critique of the judicial evaluation method. Regarding the strengths of the judicial evaluation method, Wolf (1974) made this statement:

> Perhaps the most compelling reason for using legal methdology is that it offers a useful system of evidential rules and procedures aimed at producing alternative inferences from data prior to the rendering of a judgment . . . the judical approach provides for the structured consideration of alternative arguments and inferences to keep the evaluation both intellectually honest and fair. (p. 185)

The judicial evaluation method has a number of other strengths as well. It represents a methodology for generating evaluative knowledge that explicitly relies on interpretive accounts or arguments regarding the subject matter in question. By requiring that alternative arguments regarding policy issues be developed, the judicial method encourages the piecing together of interpretations that telling a story about particular policy issues in a coherent and intelligible manner. Thus, when the judical method is used in decision-making contexts about education innovations or improvement, there is a high probability that the arguments brought to bear will be process oriented and explanatory in nature. For example, the case for acceptance of a particular innovation would have to contain explanations regarding why the innovation can be expected to operate in a particular way as well as why its operation should be judged as desirable. In short, the judicial method attempts to ensure that relevant decision makers receive recommendations based on coherent and usually detailed arguments relevant to the policy or decision in question. This stands in contrast to the frequently found situation wherein decision makers are presented with information limited to a data summary obtained for performance criteria.

In its heavy reliance on interpretivist methodology, the judicial method appears to accept the former's assumptions regarding the interrelationships between human beings and subjects, and social meaning and reality. That is to say, the judicial method accepts social reality as being constituted in large part by the intersubjective meanings, beliefs, and values that individuals attach to daily circumstances, events, and practices. However, at this point, the judical method diverges from interpretivist inquiry. As noted earlier, the latter assumes that multiple realities may exist cotemporaneously. Moreover, rather than determining truth or falsity, the interpretivist evaluation

strives to increase understanding by presenting the ways in which various stakeholders perceive the object of the evaluation. The judicial method is also concerned with presenting diverse points of view. However, in applying the legal metaphor, the judical method assumes that one point of view, or one interpretive argument, can be adjudged to be more appropriate or accurate than another. Inherent in a trial by jury is the belief that the defendant is either innocent or guilty of the crime with which he or she has been charged. The reality of guilt or innocence is constituted by a number of factors, most particularly those of social meanings (e.g., murder vs. self-defense, and social convention--laws prohibiting murder and laws allowing for particular forms of self-defense). Indeed, over time, the meanings and, thus, the realities of guilt and innocence may change (e.g., innocence due to a plea of insanity may soon be extinct). Nevertheless, an underlying assumption of the trial-by-jury practice is that the defendant is either guilty or innocent, and the reality of the situation is not dependent upon the defendant's, the victim's, or the jury's interpretation of it. Moreover, the trial-by-jury concept assumes that the "truth" of the matter (the principle) is definable and in practice may best be determined by the development and evaluation of alternative cases or arguments. Similarly, the judicial method constitutes a context within which judgments regarding which interpretive arguments appear to be the most valid and correct with regard to the issue under consideration.

The Art Criticism Model of Evaluation

Educational evaluation conceived of as a variation of art criticism represents another approach to the evaluative enterprise. Like other interpretive-based models, it was formulated in response to the predominance of scientific methodology characterizing conventional educational inquiry and evaluation. Eisner (1983) drew from art criticism as a model because he believed that teaching in a classroom is ideographic in nature and unlikely to be explained or controlled by behavioral laws. Eisner's interest in educational evaluation was to produce a "heightened awareness" of context and process of education.

> I conceive the major contribution of evaluation to be a heightened awareness of the qualities of classroom life so that teachers and students can become more intelligent within it. Connoisseurship plays an important role towards this end by refining the levels of apprehension of the qualities that pervade classrooms. (p. 339)

Eisner (1979, 1983) proposed that an underrepresented form of analysis in education is the evaluator as critical "connoisseur." He (1983) defined connoisseurship as "appreciation," which in turn means "an awareness and an understanding of what one has

experienced. Such an awareness provides the basis for judgment" (p. 341). Criticism, on the other hand, is "the art of disclosure." The author argued that criticism has to do with illumination of an educational program or event. Its end is the "re-education of perception." Connoisseurship and criticism are combined in Eisner's approach to evaluation such that connoisseurship describes prerequisite qualities necessary for an evaluator to critique a program, whereas criticism is the method of vividly portraying the program and thus making public the private apprehension of the connoisseur. The art criticism model makes the full circle, then, by claiming that criticism develops connoisseurship in others (House 1980).

Eisner (1983) identified three things that an evaluator as connoisseur covers in a criticism of an educational program: (1) description, (2) interpretation, and (3) evaluation or appraisal.

The descriptive aspect of the approach consists of a personal, subjective account of the qualities that one encounters. Such description may include number and type of questions asked in a class or some impression of a classroom as it might strike a visitor. Frequently, the descriptive approach will use the first person pronoun in written presentations. Written descriptive criticism means description explicitly through the eyes of the evaluator.

Interpretation is the evaluator's effort to understand the meaning and significance that a program has for those actors in the setting. Eisner (1983) described it as "an ability to participate empathically in the life of another" (p. 343). He suggested that varieties of social science methods might inform an evaluator of actors' understanding of a program and its meanings for various groups of actors.

According to Eisner (1983), the evaluative aspect of the connoisseurship-criticism approach is what sets it apart from ethnography, psychology, or sociology. It is meant to inquire into the educational input or value of a program. It requires the ability (and background) to consider the alternatives that might have been employed. As the third step in Eisner's approach, evaluation represents an evaluator's conclusions about the character of an educational program, its practices, and its improvement, based on what he or she has seen and has interpreted.

Candidates for critical attention, according to Eisner (1983), have to do with questions about the qualities of educational activity.

> The character of the discourse within the classroom is another candidate for critical attention. How do the children participate? What is the quality of what they

> and the teacher have to say? To what extent do they participate both psychologically and verbally in what transpires? Is what they are learning worth their time and effort? And just what are they learning? (p. 344)

Eisner (1983) asserted that questions like these reflect a conception of what is educationally valuable. Moreover, he saw no need to apologize for asking value questions in an educational evaluation. As he put it, "Only a fool would choose to attend to the trivial" (p. 345).

In summary, the evaluator as connoisseur or critic writes for consumers and other connoisseurs. Not only do they find out if a program is a good one for the consumers, but their sensitivities to subtle qualities of a program are heightened as well. There is no explicit methodology but critical review; therefore, it is the connoiseurship "stance" of the evaluator and the critical written disclosure that guides the approach. Immersion in and familiarity with a program are essential. Proper training and experience are vital. In the end, with his or her sensitivities, the experienced connoisseur renders judgment in such a way as to inform and evaluate those who are less sensitive (House 1980).

Critique of the art criticism model. The connoisseurship-criticism model of educational evaluation espoused by Eisner is another form of the "qualitative" response to conventional evaluation. The approach relies almost exclusively on the evaluator as the primary evaluation instrument. The evaluator is expected to be an informed connoisseur who renders an enlightened and enlightening criticism of an educational program.

At first glance, it seems that the connoisseurship-criticism approach might have appeal for and from a critical theory basis for evaluation. However, on close inspection, the connoisseurship-criticism approach pales when judged on its philosophical and epistomological merits. It is an approach that presents more of a technique than a profound theoretical basis for evaluation.

Eisner (1983) presented a case for a connoisseurship-criticism approach by pointing out "deleterious" consequences resulting from scientific and technological approaches to evaluation. These were the search for lawlike generalizations and assignment of certain ideas to symbols (language), a future orientation at the expense of the present, an attempt to objectify knowledge, and an orientation to the achievement of common goals and information.

Eisner's responses to these consequences of conventional evaluation might be characterized as a defense of artistic imagination. He used the nature of the arts as a contrasting basis

upon which to develop a concept of evaluation. Thus, he emphasized that educational programs should be judged on the basis of their similarities to artistic activity.

Some of Eisner's key assumptions based on the nature of art follow. First, symbolic forms directly present an idea or image that resides within the media (e.g., meaning is inherently related to the medium of its expression). Second, understanding the present experience (image) is crucial because the individuality of experience precludes any expectation of uniform practice in the future. Third, characterizations of feelings, aims, and aspirations help situate the present, and, finally, movement toward the "cultivation of production idiosyncracy" can be accepted as a worthwhile goal. Although these assumptions provide insight into the orientation of the connoisseurship-criticism approach, they nevertheless do not constitute an epistemological alternative to positivistic evaluation. However, if one accepts the assumption that "art" is an adequate analogy to education, the art criticism model may provide an alternative technique in the conduct of educational evaluation.

A major problem with the connoisseurship-criticism approach is that its reliance on an informed connoisseur as instrument and analysis tool introduces a threat of solipsism. The heavy emphasis on the perception of the evaluator (particularly in the descriptive and evaluative dimension of the approach) causes the validity of the evaluative criticism to be doubtful. Eisner (1983) argued that "structural corroboration" and "referential adequacy" are satisfactory safeguards to threats on the trustworthiness of a connoisseurship-critisism approach. However, the exclusion of other perspectives, particularly those that might inform recommended "sensitized" improvements, requires a willingness to accept the critic's account as the preferred one. Moreover, accepting the critic as expert frequently results in stakeholders' acceptance of the critic's evaluative criteria and standards, even when these are implicit (House 1980). Indeed, Eisner concluded that the use of a connoisseur-critic demands acceptance of the evaluator as sole and primary instrument. Without some greater understanding of how art as an epistomological base informs the expert connoisseur evaluator, one is left wondering how to conceive, design, execute, analyze, and conclude an evaluation. Eisner's reliance on experience does not satisfy the problem of whether one can appreciate a connoisseurship-critical evaluation as having a "better" or "satisfying" understanding of a program. The dilemma is similar to that in other "qualitative" evaluation models that apparently rely on an interpretivist philosophy of science. Without some mechanism for mediation among live perceptions of or perspectives on an educational program, the reliance on the "expert's" interpretation alone is neither warranted nor fully justified within the art criticism model.

The Professional Review Model

Professional reviews or accreditation studies of professional associations and organizations have grown in number as demands for public accountability have increased. The basic driving force of the accreditation model is succintly summarized by House (1980):

> In all these evaluations the source of standards and criteria are the professionals themselves. It is assumed that they are best qualified to judge professional merit. Procedures vary but the evaluation culminates in a holistic assessment of a professional program by other professionals. (p. 36)

The specific forms that professional reviews may assume are numerous and varied. In some cases, an outside accreditating agency conducts the review. In other cases, the review is sponsored by an organization's administration and then conducted internally by the organization's staff. An example of this type was University of Illinois' Council on Program Evaluation (COPE), which evaluated all of the academic units on the Urbana campus. Sometimes the review process is some combination of these so that an internal self-evaluation is first conducted, followed by an outside agency's review of its results and conclusions.

The professional review model is not logically tied to any particular methodology. However, some frequently used review strategies include (1) the development of evaluative criteria to focus the review process, (2) the use of checklists by review teams in rating or scoring the organization's performance in specified areas, (3) review of background documents that explicate the organization's major goals and functions, (4) review of data that are relevant to assessing the organization's effectiveness or to describing the internal environment and resources of the organization, and (5) interviews with professional staff and other stakeholders (e.g., service recipients).

Some form of the professional review or accreditation model is generally used in vocational education to comply with the Federal requirement that each local vocational education program be reviewed at least once every 5 years.

Critique of the professional review model. House included the professional review model with the subjectivist models of evaluation because the evaluative criteria used and the evaluative conclusions reached are based on the personal opinion of professionals rather than some more scientific basis. Indeed, the major critique of the professional review model is directed at its reliance on expert opinion to the exclusion of the public's or layperson's input. In this case, however, the expert is a member of the community being reviewed rather than a disinterested individual steeped in evaluation methodology. As House

noted, the reliance on professionals to police themselves has been giving way to the belief that other stakeholders should be involved in evaluation activities. This sentiment is, for example, reflected in the evaluation responsibilities that Federal legislation now gives to State advisory councils of vocational education. The point is that, with the growing concern on the part of the public that social and educational programs be held accountable for specific responsibilities, there has been a corresponding move in the direction of the public's demand to be involved in the evaluation and monitoring process.

Summary Comments Regarding the Objectivist and Interpretivist Models of Evaluation

Although there are a number of other criticisms, both positive and negative, that could be aimed at each of the evaluation models, we have limited our discussion to those relevant to the agenda of issues described in chapter 3. Here, we simply want to generalize remarks about the models.

If you summarize the differences between the objectivist and subjectivist models of evaluation, you might well develop a chart similar to this one:

Objectivist (Positivist) Categories of Inquiry	Interpretivist Categories of Inquiry
Object	Subject
Reality	Meaning
Objectivity	Subjectivity
Scientific explanation	Understanding
Facts	Values

This chart illustrates the interpretivist models' tendency to reinforce the existence of categorical distinctions rather than eliminate or override them. Our critique of each of the interpretivist models indicates that, in fact, they commit the same conceptual mistakes as do the objectivist models, insofar as they adhere to one category of inquiry to the exclusion of its opposite. Thus, for example, the objectivists make the mistake of conceptualizing human beings as totally passive objects who are acted upon by external, causal forces. The interpretivists go to the other extreme in perceiving humans totally as subjects of their experiences and activities. This view tends to come up short when attempting to explain any but the most consciously intended human activities.

Thus, a framework is needed to differentiate between when humans are indeed acting as subjects and self-forming beings and when they are reacting to influences in an "objectlike" manner. Moreover, we need a framework that comprehends social reality as being constituted by shared meanings and practices while still acknowledging that social reality can also assume an objective

status. Such a framework would help us make sense, for example, of the way in which a bureaucracy evolves from the social interaction of the individuals who make it up. Yet, it could also explain how the bureaucracy can impact upon and considerably influence the lives of some individuals who are ignorant of its existence and for whom its social character holds no meaning.

This raises another point. A great deal of time and effort has gone into criticizing the positivist program. Yet, positivism has endured and still predominates much of social inquiry and evaluation. As House (1980) and Taylor (1982) suggested, the staying power of positivism may be the promise it appears to hold for the generation of certain or more certain knowledge. Now, philosophers of science are in general agreement that no knowledge is ever really certain; indeed, the growth of scentific knowledge is now looked upon as an accumulation of propositions, theories, and so forth, which seem to be the "best" contenders for representing true or certain knowledge. However, some interpretivists reject not only the notion of certain knowledge but also the notion that any one theory or interpretation of reality can be said to be "more true" than another. This position raises a number of problems, but perhaps the most serious one is that it comes perilously close to utter relativism. And, as noted earlier, relativism is not only an unattractive alternative (e.g., we can no longer talk about mistakes due to knowledge or gaps in our knowledge), but it is also untenable because it is self-refuting. Instead of assuming a relativist position, the interpretivists would do better to provide an account of social reality, acknowledging its changing nature (e.g., social reality is historical in nature, yet retains the ideal that at any one point in time there is an interpretation that can be identified as being more accurate and normatively appropriate than others).

To summarize briefly, we have in this chapter examined a number of the major evaluation models that define the field of evaluation. Moreover, critiques of these models indicated that interpretivist or subjectivist models are generally as inadequate in some ways in their account of what constitutes evaluative knowledge as are those models called objectivist. The obvious question we need to ask at this point is where do we go from here? We think that a promising alternative to both the objectivist and subjectivist or interpretivist models of evaluation can be found in an area of inquiry called "critical theory." Although critical theory has been around the European scene for quite some time, it has only recently begun to attract attention in the United States. One of its major attractions as an alternative framework for social inquiry is its ability to account for and integrate coherently various knowledge models as well as various interests in the generation of knowledge. To acquaint those who may not be familiar with the basic concepts and logic, chapter 5 provides an overview of the theoretical basis of critical theory. Chapter 6 then provides a broad discussion of how critical theory applies to the evaluation of vocational education.

CHAPTER 5

THEORETICAL BASIS FOR CRITICAL INQUIRY

One alternative theoretical base receiving increased attention in educational research and evaluation is "critical theory." Proponents argue that critical theory effectively addresses the problems and limitations identified with positivistic and interpretivist methodologies while providing a substantive theoretical foundation of its own.

Critical theory had its formal origins in the foundation of the Frankfurt Institute of Social Research in the early 1920s. The institute's members were an eclectic group of scholars from various disciplines, most of whom were steeped in the German philosophic traditions of phenomenology and idealism. One of the major concerns of the Frankfurt school was to explain the failure of Marx's prediction of a proletariat-led revolution in the West. Their attempts to come to grips with the theoretical inadequacy often led them to make significant shifts from Marxist economic theory to other areas such as psychoanalysis, aesthetics, and communication theories. However diverse these scholars were in their approaches, they were unified in their opposition to positivism, and the latter's claims to be the sole method of acquiring objective knowledge about man and society.

More recently, interpreters of critical theory have tended to focus on the works of Habermas as the foremost representative of "modern" critical theory.* Throughout his career, Habermas has been (and still is) involved in the monumental task of developing the philosophic and epistemological foundations for a theory of society and social evolution. (Refer, for example, to his 1970a, 1970b, 1971, 1973a, 1973b, 1975, and 1979.) This enterprise has led him to an extensive number of disciplines and areas of scholarship, all of which are reflected in his writings.

Although critical of positivism as a theory of knowledge, Habermas does not view interpretivism or naturalistic inquiry as a viable alternative to the positivistic program. Rather, he attempts to provide a theory of knowledge incorporating the positive aspects of positivism and interpretivism and refuting their inadequate assumptions. We will not attempt to discuss the full body of Habermas's work in this manuscript. Instead, we will discuss those aspects of his work most relevant and helpful to our explication of a new framework for vocational education evaluation.

---

*Throughout this discussion, we will be explicating from the collection of Habermas's work rather than from any one single source.

Human Action, Interests, and Knowledge

Habermas viewed history as the expression of the activities of human subjects over time. History can then be viewed either from the perspective of the individual's biography or from the perspective of the development or evolution of society. Historically speaking, social evolution is constituted by and expressed in the dialectical interplay between two forms and domains of human activity. The first is activity within the domain of material nature (i.e., that realm of being that lies outside of and independent of human existence and consciousness). The point of human activity in this domain is, has been, and always will be to exert control over and to subjugate nature to human purposes. Thus, activity in this realm is *instrumental* in nature.

During the course of pursuing instrumental activity, human subjects also conduct activities in a second domain--that of *practical* action. Through practical activity, human subjects create, generate, and maintain forms of symbolic communication as reflected in the meanings, norms, rules, institutions, and so forth that characterize a society. Through the creation and maintenance of symbolic communication, social integration and survival are achieved.

This notion of history is the context within which reality and our knowledge of reality are constituted. Moreover, this view of history is similar to that of Marx.

> Social reality is neither something wholly "outside" of the subject nor is it simply a creation of human thought. Rather reality is conceived as formed and constructed through practice and labour. Through practice and labour the human species synthesized and alters the material world and thereby transforms nature *qua* known as well as itself. The objects of human perception are themselves the products of the self-generative and self-formative activity of the species. What we understand by nature or the human species changes over time as both are actively transformed. The process of knowing cannot be separated from historical being. (Held 1980, p. 190)

For Habermas as well as Marx, human knowledge is inextricably linked to human activity as it is played out over time in the domains of technology, symbolic communication, and, ultimately, social evolution.

Viewed in this way, human knowledge is generated and formed by human interests. Habermas (1971) argued that the "technical cognitive interest" constitutes knowledge enabling human beings to shape nature in conformity to their needs and wants. The "practical cognitive interest" constitutes knowledge whereby

social integration is maintained. Integration is achieved because practical knowledge also "facilitates the understanding and the reaching of understandings with others" (Bredo and Feinberg 1982, p. 272). Although the technical and practical interests and the knowledge constituted by them can be conceptualized as being separate and distinct, they are in fact dialectically interrelated. Thus, the development of technology may be impeded and even denied by certain forms of social and political communication and organization. Conversely, advances in human technological knowledge may effect changes in social communication that are inconsistent with individuals as human subjects. The dialectical relations between the technical and practical interests constitute a third human interest that Habermas called the "emancipatory interest." The emancipatory interest generates and constitutes knowledge that seeks to resolve the inconsistencies and contradictions between technical and practical knowledge and thus emancipates humans from the constraining conditions of either or both.

Habermas argued that the form and content of the knowledge constituted by each of the three cognitive interests provides a basis for the form and content of three separate and distinct modes of inquiry. Thus, technical knowledge is the focus of the positivistic sciences such as physics and astronomy in the natural realm and psychology and economics in the social scientific disciplines. In the positivistic sciences, what counts as knowledge are the deductive-nomological systems of propositions that assert lawlike regularities between measurable and observable entities and that have been confirmed through testing over time. In the social sciences, the search for technical knowledge has been exemplified by the search for behavioral regularities that are ultimately explained by reference to psychological cause and behavioral effects.

Practical knowledge has been the object of inquiry for the hermeneutic sciences such as history and anthropology. The "objects" of inquiry in these sciences are the meaning complexes intersubjectively shared by societal members at any one point in time. The methodology of acquiring practical knowledge consists of developing interpretations of the shared meanings and symbolic interactions. These then are expected to provide understanding regarding the point of human activities within a particular historico-social situation. In contrast to the positivistic sciences' emphasis on causes, the hermeneutic sciences seek to understand the reasons and motivations underlying human action by reference to the individual's own interpretation of social life and meaning.

The emancipatory interest constituting the domain of the critical sciences generates knowledge enabling individuals and society to be liberated from systematic distortions in communication and understanding. Such distortions obscure individuals' awareness of themselves as human subjects and their historical positions.

> While the empirical-analytic sciences are structured to generate knowledge in the form of laws and theories which can account for regularities in 'observable' phenomena and the hermeneutic sciences are structured to allow the recovery of a particular meaning of an action or expression, the critical sciences, Habermas maintains, unite an interest in nomological and interpretive knowledge aimed at facilitating the process of self-reflection. (Held 1980, p. 295)

Central to understanding the need for self-reflection is an understanding of ideology and its role in the history of society.

Ideologies and Social Evolution

Ideologies are the systems of intersubjective beliefs, meanings, and values within which social institutions are generated, maintained, and changed. They serve to justify and legitimate the existing institutions, norms, and rules characterizing the practical realm of human knowledge and activity. In doing so, they provide a coherent and integrated interpretation of society's technical and practical domains and the individual's position or status within society. Thus, ideologies are the "cement" that maintains social integration. In terms of social evolution and survival, this is an essential function since societies are continually faced with potential disruption and disintegration.

The potential for social disintegration is constant for the following reasons. Like all creatures, human subjects have various and diverse interests (i.e., wants and needs). However, the ability to satisfy all wants and needs is sometimes limited by natural resources and the knowledge required to employ these resources. As a result of the scarcity of resources and technical knowledge, not all wants and needs of society members will or can be fulfilled simultaneously.

Moreover, whereas the interests of some society members will go unfulfilled, deprivation of other interests will be class or group specific. Social institutions along with their constituent norms, rules, and values are the means by which society represses human interest (at least) sufficiently to allow for social survival. Moreover, social institutions are the means by which the fulfillment of some interests and the deprivation of others are distributed among groups of individuals within the society.

As noted earlier, Habermas conceived social evolution to be the dialectical interplay over time of developments within the technical and practical domains of human activity. This is manifested in history as changes in the nature of human interests, human knowledge, and social institutions and patterns of organization. As Held (1980) pointed out, Habermas agreed with Freud.

> The effectiveness of social institutions in facilitating survival occurs at the cost of repression of wants and needs . . . the level of necessary social repression was a function of the level of development of productive forces; as the technical power of a society to control the forces of outer nature increases, the constraints of scarcity are progressively overcome, thus decreasing the degree of socially necessary repression. As the level of repression diminishes, the institutional framework of a society can be changed to accommodate higher levels of needs gratification. (p. 277)

Although developments in technical knowledge may allow a society to fulfill new or more numerous human interests, the forms of communication and interaction constituted by the prevailing or predominant ideology will work against major changes in society's institutional patterns of organization. This tension between potentiality and actuality are the sources of distortions in social communication and interaction.

Distortions are allowed to exist because ideologies serve as frames of reference and consciousness for most individuals in society. Indeed, ideologies frame and constitute individuals' experience of social interaction and the institutional arrangements underlying that interaction. Because of this, ideologies can mask or obscure the authentic interests of individuals as human subjects. That is, an ideology motivates individuals to accept certain institutional arrangements as "given" or as serving their wants and needs even when reflection and critical analysis would indicate otherwise. Ideologies, like the social institutions embodying them, are able to maintain social integration so long as they are perceived as coherent and legitimate rationalizations for the existing order.

According to Habermas, human freedom is dependent on society's ability to develop in both the technical and practical realms while (at worst) minimizing and (at best) abolishing altogether the sources of systematic distortion in social communication and interaction. Sometimes progress in human freedom is initiated by unreflective human activity. In these cases, a society will find itself in a situation that strains the coherence and intelligibility of the prevailing ideology. In such a case, the tension between the society's technical capacities and its institutional distribution of interest fulfillment becomes severe enough that it can no longer be obscured and legitimated by the dominant ideology. Those groups who have experienced the most serious deprivation will most likely recognize the tension and push for changes in social institutions that will more accurately reflect the society's needs gratification capacity.

A simplistic example of this is the economic depression that hit the United States and other Western nations during the 1920s and 1930s. Although ultimately a result of human action and

practice, the depression was not intentionally brought about by anyone. Nevertheless, once the full effect of the depression was underway, it became obvious that some groups were being severely deprived of even basic necessities, whereas others enjoyed the luxuries that great wealth can buy. Moreover, the prevailing ideology could not adequately explain or legitimate the disruption and disintegration; institutions were changed and introduced in order to bring about a new distribution of the Nation's wealth.

Thus, one way ideology is exposed is via society's drifting into situations that strain the legitimacy of prevailing modes of social communication and understanding. However, ideology can also be exposed through the conduct of the critical analysis of society wherein a reconstructed interpretation of society both identifies and explains the operation of ideology in social life and interaction. Insofar as individuals provide social communication with its meaning and objective status in reality, they can shape and mold it to satisfy their needs and desires. Moreover, this shaping and molding can be done in an ideological or uncritical and unreflective manner or according to what critical reason dictates about the quality of human life. For Habermas and other critical theorists, the latter type of activity is called praxis. As Jay (1973) explained:

> Loosely defined, praxis was used to designate a kind of self-creating action, which differed from the externally motivated behavior produced by forces outside man's control . . . one of the earmarks of praxis as opposed to mere action was its being informed by theoretical considerations. (p. 4)

Next, we discuss how one goes about identifying ideology and its sources.

The Ideal Speech Situation

In order (1) to observe distorted communication and interaction and (2) to judge the validity of particular interpretations of social life, the critical theorist must have some notion of what would constitute undistorted communication and understanding. Habermas (1973) attempted to develop a theory of communicative competency or "universal pragmatics" that provided the foundations for characterizing such a notion. In this theory of communicative competency, Habermas focused on the universal functions of language as the predominant medium for human and social communication. He argued that each and every speech act fulfilled at least three separate but interrelated functions. The first was conveying or transmiting some content about some state of affairs. The second function was positing or establishing a social relationship between speaker and audience. The third which is derivative of the first two functions, is expressing

some intent on the part of the speaker in repeating what he or she is saying.

In addition to fulfilling these functions, each and every comprehensible speech act presupposes that validity claims related to these functions have been met. Thus, as Bredo and Feinberg (1982) explained:

> When one says something in a particular situation and before a particular audience one in effect claims: (1) that the propositional content of what is said is true; (2) that one has the right to adopt the social position that one adopts (i.e., that one has the right to speak in this way in this situation); and (3) that the expressive aspects of the utterance are genuine (i.e., that one is neither self-deceived nor deceiving of others). These three validity claims, then, refer to the truth, legitimacy and veracity of the utterance and its speaker. (p. 282)

In everyday affairs, it is usually the case that the validity claims of speech acts go unscrutinized and unquestioned. There is an "accepted consensus" that communication is validly fulfilling the three universal functions. This is referred to as "first-order" communication or "communicative interaction." However, the accepted consensus may break down when one or the other or all of the validity claims underlying first-order communication are called into question. When this happens, "second-order" communication must occur to settle validity disputes or issues. The activity of conducting second-order communication is referred to by Habermas as "discourse." Resolution through discourse of issues regarding the validity claims of communication rests on a "rational consensus" as opposed to the accepted consensus of everyday speech acts. The difference lies in the fact that rational consensus is arrived at through the reflective and critical examination of validity claims, whereas accepted consensus rests on unreflective, hidden, and simply taken-for-granted presuppositions.

Inherent in Habermas's conceptualization of rational discourse is a characterization of the ideal speech situation. That is to say, in order for valid or authentic discourse to occur, the following minimal conditions must be met in the speech or communication context:

- Any and all forms of coercion or domination must be absent from the context.
- There must be freedom of access or entry to the context of discourse to all individuals.

- There must be no exclusion from consideration within the discourse context of any aspects of the communicative activity.

Because these conditions are inherent in Habermas's conceptualization of rational discourse, they are also inherently related to and presupposed by first-order communication.

Ideally, the participants within a discursive context can be said to be "communicatively competent." Bowers (1984) noted that "communicative competence . . . can be understood most simply in terms of the individual's ability to negotiate meanings and purposes instead of passively accepting the social realities defined by others" (p. 3). Beyond possessing some skills in the advancement, critique, and negotiation of arguments, communicatively competent participants must also demonstrate a willingness to play by the rules of critical discourse. "For example, discourse must not be viewed in terms of conflict or competition and the outcome must not be seen in terms of winning or losing" (Cherryholmes 1982, p. 65). Moreover, "in a discourse, the 'force' of the argument is the only permissible compulsion, whereas the co-operative search for truth is the only permissible compulsion, whereas the co-operative search for truth is the only permissible motive" (Habermas 1973, p. 168). Indeed, it is getting to the truth of the matter that enables participants in discourse to lay aside their particular interests in the general and public interest of generating valid knowledge.

Finally, it should be noted that critical discourse does not have to be face-to-face. As Cherryholmes (1982) pointed out, "It may also be thought of as a metaphor where the pursuit of truth is conceptualized as an ideal face-to-face interaction" (p. 66). Thus, discourse may assume various media of expression just so long as the three conditions noted earlier are met. Examples of such media include professional journals or newsletters, newspaper articles, television spots on public access channels, or any combination of these.

Habermas went on to argue that to date the conditions of the ideal speech situation have not been met in most societies due to the operation of ideology. Ideology introduces the elements of coercion and domination into first-order communication by exercising a hold on individuals' sense of who they are and what it is they need or want. In other words, although ideologies are constituted by individuals' subjective belief in and adherence to the former's interpretations of reality, they assume an external relation of domination over individuals by framing their categories of consciousness, experience, and meaning.

Moreover, the functioning of ideology results in an inequality of access to the arenas wherein material and symbolic distributive issues are determined. Indeed, this denial of access to some groups is rationalized and legitimated by the prevailing

ideology. Along these lines is Lowi's (1969) contention that the current ideology in the United States legitimates the fact that only those groups whose members have resources enough to pursue their interests are allowed access to political arenas in which policies of social significance are determined.

Finally, insofar as it provides the categories and structure of meanings through which social experience is mediated, an ideology will restrain the realm of what exists and what is possible. In doing so, it demarcates what it is that can reasonably be discussed. To quote Bowers (1982) again, the "political nature of language" is such that "the elimination of words can lead to areas of existential silence and political impotence in terms of communicating and negotiating with others over issues affecting lived but inarticulated dimensions of life" (p. 538). In short, we do not tend to scrutinize and discuss what is not perceived to exist.

Held (1980) succinctly summarized Habermas's position:

> It is Habermas' contention that in every communicative situation in which a consensus is established under coercion or under other similar types of condition, we are likely to be confronting instances of systematically distorted communication. This is, in his view,the contemporary formulation of ideology. On this account ideology is, as Trent Schroyer put it, "those belief systems which can maintain their legitimacy despite the fact that they could not be validated if subjected to rational discourse." (p. 290)

The centrality of criticism and discourse in Habermas's work becomes obvious at this point. It is the only process by way of which undistorted and noncoercive understandings about the individual, society, and social interrelations can emerge. As Cherryholmes (1982) pointed out:

> Criticism is needed to disclose and peel back the layers of value and commitment embedded in interpretations and explanations. Criticism does not lead to value free knowledge claims but instead surfaces, analyzes, and scrutinizes existing values with the goal of emancipating individuals from social institutions that are historically and culturally conditioned. (p. 61)

The Methodology of Critical Theory

As one author pointedly stated, "Unlike the positivist tradition which has a sophisticated system for translating theory to methodology, critical theory gives little direction on how to's" (Mathison 1982, p. 2). Nevertheless, beyond specifics, critical

theory does provide some broad methodological direction. However, let us first look at how Habermas and other critical theorists perceive the methodological stances of positivism and interpretivism. Habermas and others claimed that both positivism and interpretivism are partial and one-sided.

Their respective partiality derived from the fact that each excluded a major methodological focus of the other (i.e., positivism excludes from consideration the subjective components of human action, whereas interpretivism excludes from consideration an analysis of causal-like relations that are independent of individuals' consciousness). In contrast to either of these positions, critical theory accords both cause and intentionality an important role in accounts of human and social life. Habermas pointed to Freudian psychoanalysis as one critical science exemplifying this principle.

Psychoanalysis as a Critical Science

According to Habermas, the point of psychoanalysis was to enable the individual to penetrate the surface appearances of his or her life in order to reveal the deeper reality that guides behavior. Through the therapeutic process, the individual comes to see that repetitive patterns in their actions are not, as is often believed, the result of consciously made life choices, but rather are the result of unconscious motives and desires. The latter induce the person to act upon life choices that are neurotic in nature and that inhibit the individual's well-being. The unconscious motives remain obscured to the individual because they are repressed over a lifetime and do not show themselves in their true form. Rather, they work through the cover of interpretations of situations and choices that are acceptable to both the individual and society. Thus, individuals who have not undergone psychoanalysis live their lives out by uncritically accepting as authentic their conscious intentions and choices even though the latter may actually be a result of repressed motives and drives. Those individuals who enter therapy are enabled over time to reflect on the true causes of their actions and in so doing can liberate themselves from negative impacts.

Thus, on this account, the analyst as investigator will need to consider the patient's conscious intentions and life interpretations since the latter are the media through which the patient's unconscious is manifested. However, the analyst will also have to seek out and identify behavioral regularities indicating those unconscious causes that are operative in the patient's psyche. Exclusive attention to conscious intentions or underlying causative motives is not adequate for the analytic process; both must be addressed to assist the patient in learning how to live a nonneurotic life.

Before discussing how psychoanalysis may illuminate critical social theory, some other points should be emphasized. First, it is important to note that psychoanalysis requires the patient, with the analyst's assistance, to reconstruct his or her life history. That is, the patient must begin the process of reconstructing his biography so that insights into unconscious motives and desires can lead to and be integrated into a meaningful and more intelligible interpretation of his or her conscious life history.

A second point to be emphasized has to do with the analyst's role in the psychoanalytic process. While bringing to bear his or her own expertise and knowledge of psychoanalytic theory, the analyst will nevertheless rely on the patient's interpretive and self-reflective capacities to provide information and the basics relative to reconstruction. In this sense, the analyst, like the patient, is a participant rather than an observer of the process, as both collaborate to develop a coherent reconstruction. Moreover, the reconstruction process can proceed only insofar as both patient and therapist share a mutual consensus on the truth, legitimacy, and authenticity of the resulting interpretation. When consensus on any of these aspects breaks down, therapy comes to a standstill and if not repaired, halts altogether. This stands in contrast to any notion of therapy in which the analyst/expert informs the patient what his or her problems are all about and prescribes steps to rectify those problems.

This raises a third point, which is the fact that the reconstruction process (and, thus, the psychoanalysis itself) is an educative enterprise and experience for both patient and analyst. Thus, although the analyst's experience and knowledge may serve to initiate and expand the patient's self-reflective capacities, the patient's continual insights through self-reflection will hone and extend the analyst's psychotherapeutic and interpretive capacities. In fact, neither the analyst nor the patient is able to predict what the full reconstruction of the patient's life history will look like prior to the conclusion of analysis. Thus, for both analyst and patient, the analytic process will unfold as an enlightening experience.

Finally, it should be noted that within the psychoanalytic context, the patient becomes initiated to and involved in the process of praxis (i.e., self-reflective and self-forming activity as a human subject). Thus, patients not only reconstruct their life history in analysis, but they also continually reflect on their present activities and attempt to integrate them into an expanded and more intelligible interpretation of their life situation. This process does not terminate with the conclusion of analysis. Rather, it is anticipated that the psychoanalytic experience will enable the patient to continue to lead a praxis-oriented life that will include an ongoing activity of life reconstruction and, thus, liberation from the coercive effects of the unconscious.

Using the model of psychoanalysis as a critical science, we can begin to see how Habermas's conception of ideology and of the ideal speech situation operate in a methodology for critical social analysis.

Critical Social Analysis

Just as neurosis obscures from the individual the reality underlying his or her conscious intentions and activities, ideology obscures from social members the sociopolitical and socioeconomic structures and relations that underlie their particular conventions and institutions. The gap between the hidden ideological foundation and the conscious understanding of social life gives rise to systematic distortions in social communication and understanding. Thus, critical social analysis must attend to the practical domain of human activity as it appears in the society's meaning structures, institutions, practices, and rules. However, it must also seek out patterns of behavioral regularity as cues or symptoms of underlying ideological causes. As in psychoanalysis, the critical social analyst must also become involved in the process of reconstructing an interpretation of social phenomena that will make intelligible the sources of and reasons for distorted social communication.

The concept of the ideal speech situation figures prominently into this process. As Held (1980) pointed out:

> Critical theory is . . . grounded in a normative standard that is not arbitrary, but inherent in the very structure of social action and language. It is just this anticipation of an ideal form of discourse which can be used as a normative standard for a critique of distorted communication. (p. 256)

By comparing the existing forms of social interaction with that which should occur under conditions of ideal speech, the critical analyst can begin unraveling appearance from reality in the process of developing a rational reconstruction of social life.

Knowledge generated through critical social analysis is rendered in the form of interpretations (i.e., rational reconstructions) of social life that are explanatory in nature. At this point, the question arises, Upon what basis does one determine whether to accept a particular interpretation as true, accurate and valid? The response to this will seem facile to some and inevitable to others. Decisions regarding the validity and appropriateness of a particular reconstruction can only be formed within the context of rational discourse. Bredo and Feinberg (1982) posed and responded to the question:

> Are there then any criteria for deciding when a validity claim has been met--for example, for deciding when

> a statement is true? In this view there are none outside of the process of argumentation itself, where this presumably involves communicatively competent persons. (p. 283)

However, discourse itself is not a single, isolated act. Rather, it is an activity that persons participate in over time. Thus, decisions regarding the validity of a particular reconstruction or regarding the relative strengths of alternative reconstructions are not formed over night. Instead, they <u>evolve</u> out of the discursive context.

Moreover, as in the reconstructed biographies of psychoanalytic patients, a reconstruction of social life can never be accepted as valid in a final sense. Rather, it will require continuous reworking and extension to account for changes in historical conditions. In this sense, it will serve an educative function for both the critical theorist and his or her audience in their conduct of praxis over time. It is the process of critical interpretation, coextensive with the ongoing activity of discourse, which generates knowledge that allows individuals to be self-forming subjects.

CHAPTER 6

CRITICAL INQUIRY AS AN ALTERNATIVE FRAMEWORK FOR VOCATIONAL EDUCATION EVALUATION

Earlier in this manuscript, we suggest that an evaluative framework provides a view of the subject matter of evaluation as well as an account of how that subject matter is to be appropriately treated. In doing this, the framework establishes criteria of relevance to identify the essential dimensions of the subject matter. Moreover, an evaluative framework will "secrete" (to borrow Taylor's 1974 term) evaluative criteria whereby evaluative judgments regarding the subject matter can be formed.

In this chapter, we will indicate how critical theory or inquiry can serve as a basis for an evaluative framework that stands as an alternative to the technological framework. Although definitely not a "how-to" chapter, it does attempt to provide readers with a feel for what critical evaluation might be like.

View and Treatment of the Subject Matter

The technological framework provides an economic view of the subject matter of vocational education. It does this by justifying vocational education on grounds of economic rationality and by applying the industrial production metaphor to the vocational education process. In this view, vocational education can be treated as a separate and distinct process independent of the context within which it operates. Similarly, the component elements of the vocational education process (i.e., inputs and outputs) can be treated separately and distinctly for purposes of evaluation.

In sharp contrast to the technological framework, a critical evaluation framework provides an anthropological view of the subject matter. As used here, the term anthropological means the network of roles, practices, and institutions that characterize a particular community or group of individuals over time. On this view then, the relevant metaphor is the biological organism. (The root source of the term "culture" was used in reference to cultivating living things or organisms.) In biology, the organism is studied in holistic terms and is conceived to be something more than just the sum of its parts. Biological understanding of an organism includes both an understanding of how the parts are related to one another and to the whole and an understanding of the organism's evolutionary path. Indeed, a thorough knowledge of the nature of an organism requires an understanding of how it has come to be what it is. In the social realm, the biological organism is a social culture. Evolutionary understanding becomes historical understanding. Study of social culture must proceed

by acquiring an understanding of how the cultural parts (i.e., roles, practices, institutions, meanings) relate to one another and to the whole. Understanding a sociocultural organism must take into account the organism's attempts to adapt to changes in both its internal (i.e., social-cultural) environment and its external (i.e., natural) environment.

In this view, then, vocational education as a subject matter is to be viewed as a cultural institution. Moreover, as a subject matter, it can be viewed only in relation to (1) the sociocultural institution of education generally and (2) the relation of the latter to the more general culture within which it operates. Sirotnik and Oakes (1984) proposed a statement to this effect:

> The methodology of critical reflection demands that participants attend to how educational structures, content, and processes are linked to the social and political forces in the setting and to the larger social, political, and economic context in which the school is situated. (p. 63)

The treatment of this cultural subject matter would, of course, be a critical one. That is to say that evaluation inquiry would proceed through development of interpretive or reconstructive explanations of educational and vocational education phenomena. Such explanations would continuously be subjected to, in Sirotnik's (1984) words, "rigorous and sustained dialogue" within a discursive context.

In the technological evaluation framework, the natural locus for the role and function of vocational education is training students in the skills needed by business and industry. A cultural and critical treatment of vocational education changes this role and function considerably. Within a critical framework, vocational education's primary role and function becomes interrelated to that of education in general. Thus, the appropriate role and function of vocational education and of education generally, is that of providing and generating technical, practical, and critical knowledge and skills individuals need to become reflective and critical (i.e., emancipated) sociocultural members.

Critical evaluation thus becomes a mode of evaluation inquiry aimed at the development of valid explanations of the vocational education setting so that problems with or constraints on efforts to accomplish this broader function can be dealt with and overcome. Moreover, in talking about "hermeneutic policy analysis," which shares the same characteristics of critical evaluation, Dryzek (1982) defined it "as the evaluation of existing conditions and the exploration of alternatives to them, in terms of criteria derived from an understanding of possible better

conditions, through an interchange between the frames of reference of analysts and actors" (p. 222).

In the following sections, we will further explain critical evaluation. We will examine how, within a context for generating evaluative knowledge, the ongoing activities of reflection and discourse are used for framing evaluation problems, interpreting problem dynamics, and overcoming problems. We will then discuss a number of issues related to the practical application of a critical framework in the vocational education context.

Context for Generating Evaluation Knowledge

As noted in chapter 5, critical or theoretical discourse picks up where practical discourse leaves off. More specifically, it picks up when any of the four validity claims underlying communication are brought into question. Remember that these four conditions consist of the following characteristics: (1) the speech act is comprehensible and intelligible, (2) its propositional content is true, (3) its utterance is a sincere or authentic act, and (4) it reflects a legitimate relationship between speaker and audience. Thus, when accepted consensus regarding the intelligibility, veracity, and authenticity or legitimacy of speech or knowledge claims breaks down, the stage is set for the initiation of critical discourse. Some examples of hypothetical questions that may initiate the need for critical discourse include the following:

- What are the short- and long-term consequences of educational tracking practices, and whose interests do they serve in what manner?
- What are and should be the major functions served by vocational education at the secondary level, and at the postsecondary level?
- To what extent does vocational education contribute locally and nationally to the perpetuation and elimination of economic and social discrimination against special needs groups (e.g., ethnic minorities and the handicapped)?
- Do current vocational education curricula socialize students into unreflective acceptance of business-dominated reward and value structures?

Questions such as these are likely to emerge during self-conscious and reflective efforts at improving the quality and conditions of the educational experience. The fact that existing and taken-for-granted beliefs and knowledge are insufficient to address them indicates the need for a school-based context for generating evaluative knowledge.

Figure 1 displays the core elements of a critical evaluation context. These elements should not be viewed as methodological components implemented in a linear fashion; rather, they are consitutent "moments" or phases that collectively establish the context within which critical discourse and evaluation may be conducted.

Framing Evaluation Problems

One significant phase of a critical evaluation is the framing of the evaluation problem(s) or question(s). The notion of framing educational problems includes both the identification of problems to be addressed during the evaluation and the interpretations that stakeholders have of these problems based on their experiences.

Particular attention must be paid to the "generative themes" employed by stakeholders in experiencing and interpreting the vocational education process. Generative themes are the assumptions and presuppositions that constitute and guide individuals' views of the world and their relationship to it (Freire 1981). As such, generative themes and the interplay among them directly relate to actions that constitute and transform social and educational reality. Thus, how individual stakeholders frame their experience of an educational problem will also be related to the generative themes they employ. Understanding the cultural setting and clarifying generative themes are activities that provide the backdrop for formulating evaluation problems or questions to be addressed. Once formulated, the evaluation problems are continuously analyzed throughout the evaluation process for possible modifications (i.e., for possible reframing).

Interpreting Problem Dynamics

One of the major aims of critical evaluation is to construct valid interpretative explanations around which rational consensus can be formed about how and why the vocational education reality has evolved into what it is. Only in this way can stakeholders clearly identify problem sources and situational constraints. One of the distinctive characteristics of critical evaluation is the participatory nature of the interpretive process. Participation by stakeholders is necessary because the vocational education reality is primarily constituted by the activities and attendant beliefs and values, and meanings of individual stakeholders. Thus, authentic change (i.e., improvement) can only occur through the conscious activities of these same individuals and groups.

Participation is also necessary in order to conduct discourse through which the validity of proposed interpretive explanations can be examined. Participation in discourse helps to

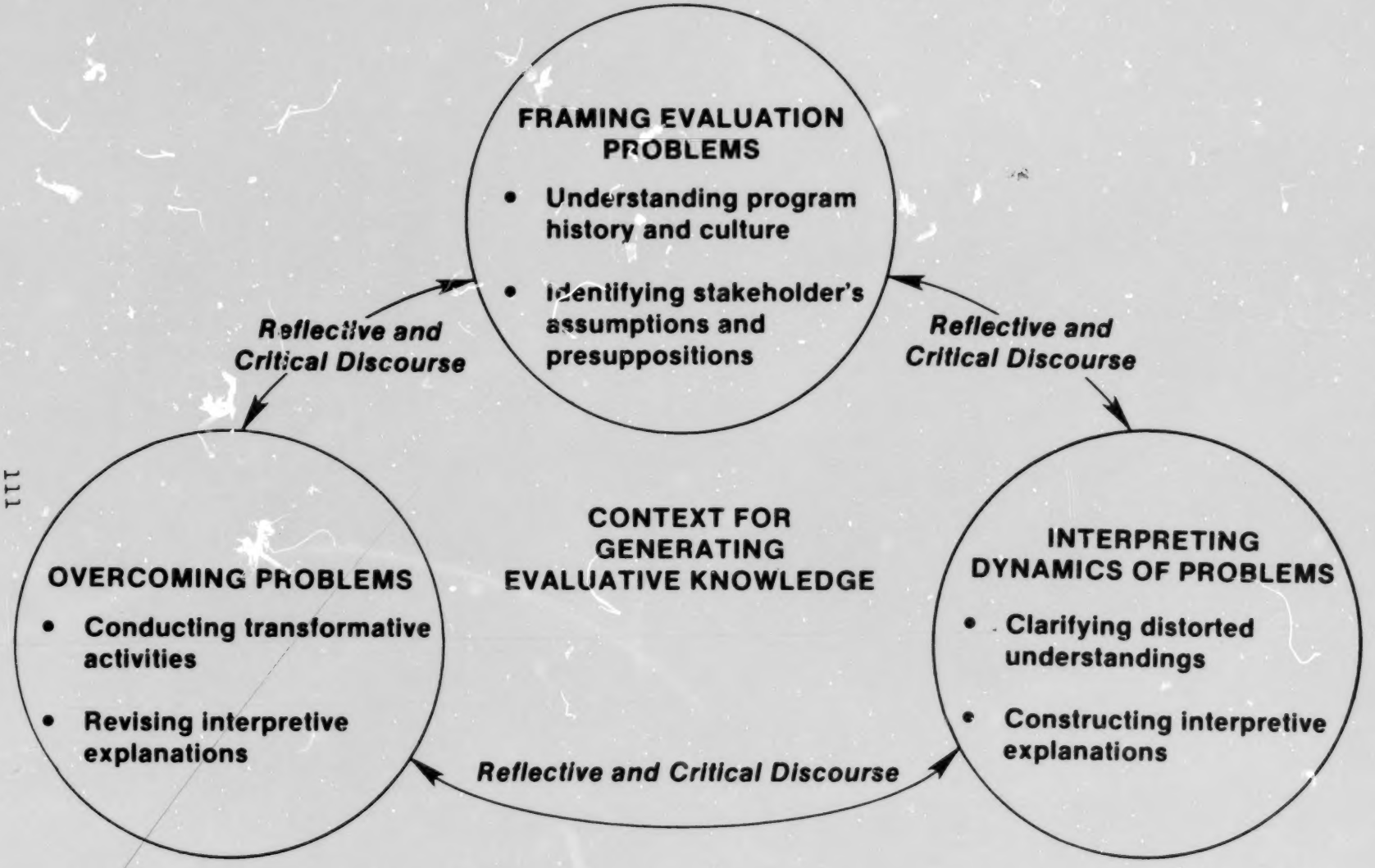
FRAMING EVALUATION PROBLEMS
• Understanding program history and culture
• Identifying stakeholder's assumptions and presuppositions
Reflective and Critical Discourse
Reflective and Critical Discourse
CONTEXT FOR GENERATING EVALUATIVE KNOWLEDGE
OVERCOMING PROBLEMS
• Conducting transformative activities
• Revising interpretive explanations
INTERPRETING DYNAMICS OF PROBLEMS
• Clarifying distorted understandings
• Constructing interpretive explanations
Reflective and Critical Discourse

Figure 1. Critical evaluation in vocational education

identify possible cases of distorted understanding on the part of stakeholders. Distorted understanding may refer to mistaken readings of the situation, inappropriate value positions, or claims being proposed that are based on self-serving interests in setting outcomes. Explanations based on distorted understanding of the vocational education reality are themselves distorted; thus, valid interpretive explanations must be sought through discourse involving the participation of relevant stakeholder groups.

The process of constructing an interpretive explanation is similar to that of deciphering an obscure code. The interpreter is confronted with an array of complex and multifaceted information from which a coherent reading must be fashioned. Thus, various methodologies for obtaining information (including empirical-analytic and interpretative ones) may be used in constructing the interpretation. Moreover, state-of-the-art research findings may useful in making sense of the phenomena being studied.

For example, vocational student placement rates are used in assessing the effectiveness of vocational education programs. If placement rates for a particular program are high, it is frequently assumed that the program is operating effectively. On the other hand, low placement rates for a particular program often are assumed to indicate that the program needs to be improved or eliminated. However, placement rate data in and of themselves are of limited value, and any conclusions made about program effectiveness based solely on these rates may be presumptuous. Until the data are described and interpreted their conclusions about program effectiveness run the risk of being invalid.

Sirotnik and Oakes (1983) suggested that data analysis is a key activity in the interpretive "decoding" process. In one study, they utilized school staff in small groups to reflect on the data so that each group could make its own descriptions and interpretations of the meaning of the data. They then concluded that "while these decoding activities are likely to be most effective with small groups working together, a constant flow of information--a dialogue--must be maintained among the whole staff" (p. 115). These sessions can be taped and shared among the groups for further decoding. During this decoding process, themes that pervade the school should emerge, which, in turn, lead to possible action. Sirotnik and Oakes (1983) also note that "these themes and the actions they imply then become the center for further inquiry and eventually of the school improvement program as well" (p. 116).

An important element of the interpretive activity is the formation among participants of a rational consensus regarding the validity and appropriateness of a particular interpretive explanation. As noted in chapter 5, a rational consensus differs from an accepted consensus insofar as the latter is constituted

by our everyday, taken-for-granted thinking, whereas the former is derived from a radical critique of these. However, the question arises regarding the nature of the grounds upon which a rational consensus must rest. The response to this question is succinctly summarized by McCarthy (1973):

> The analysis of speech shows it is oriented toward the idea of truth. The analysis of "truth" leads to a notion of a discursively achieved consensus. The analysis of "consensus" shows this concept to involve a normative dimension. The anslysis of the notion of a grounded consensus ties it to a speech situation which is free from all external and internal constraints, that is, in which the resulting consensus is due the force of the better argument. (pp. 153-154)

Thus, what is meant here by "the force of the better argument" is a key consideration in a critical evaluation. Positivist social research has often adopted the correspondence theory of truth. According to this theory, a statement may be accepted as true if and only if its terms are satisfied or "correspond" to the factual state of affairs in reality. However, as Cherryholmes (1982) observed, "This is an appealing idea. But it runs into the fundamental problems that statements are different from things in the world. Formal logic provides rules for testing statements against statements but there are only guidelines for testing statements against things" (p. 60). The point is that no science, physical nor social, can claim to generate knowledge that can be proven to be true or false with absolute certainty.

Currently, most philosophers of science recognize that knowledge is "preferred" because there is consensus among members of the relevant scientific community that it is to be preferred. Obviously, this consensual approach to truth and knowledge raises again the spectre of accepting as preferred knowledge that which is in fact biased or distorted. However, this view of consensual truth does not necessarily mean that decisions regarding what will count as knowledge will be wholly subjective and biased. This is because scientific knowledge can always be subjected to theoretical and critical discourse. Cherryholmes (1982) noted:

> Habermas developed the idea of communicative competence and discourse to deal with the problems of bias and distortion. Free and unconstrained discourse that leads to increasingly reflected levels of thought is used to test the truth of statements. This allows one set of coherent statements to be evaluated against another as well as to reveal the purposes and normative commitments embedded in them. Habermas's goal of "demystifying false consciousness" can be approached as arguments are increasingly reflected. (p. 61)

The point is that the interpretive explanations generated within a critical evaluative context will be accepted only insofar as they hold up when compared to rival interpretations and when compared to the relevant evidence. What counts as relevant evidence will itself be determined through the discursive process and may include findings or conclusions based on interpretive as well as causal modes of inquiry.

Finally, it should be noted that the critical process cannot be expected to yield a single explanation that is valid for all times and in all contexts. Rather, it can be expected to yield a valid explanation of current vocational education reality that will illuminate present problem sources and suggest present opportunities and strategies for transformative action. With the passage of time, this interpretive explanation must be subjected to reanalysis, critique, and reconstruction.

Overcoming Problems

Once the new reality is clarified, obstacles confronting vocational education improvement efforts need to be identified. Obstacles are those problems that impede change efforts and are perceived to be real by those in the vocational education setting. If identified problems are not perceived as problems by those in the setting, then attempts to resolve those problems become solo efforts by the evaluation team. As a result, participants in the setting become removed from the evaluation effort, which, in turn, loses the power of critical inquiry.

Transformative actions are attempts to penetrate and remove obstacles impeding vocational education improvement (i.e., they are efforts for change). Effective transformative actions solve problems and may cause new interpretative explanations of reality to emerge. These new interpretations illuminate more problems and, consequently, call for further actions. In essence, reality transformation is a perpetual process in which actions play a pivotal role in perceiving and resolving new problems.

Strengths of the Critical Evaluation Approach

The critical evaluation approach for vocational education has several key strengths that distinguishes it from current evaluation efforts in vocational education. A brief description of these strengths is highlighted here.

Understanding the Past

Vocational education programs do not exist outside of their historical context. A historical understanding of the ways the vocational education program is based on the community and school

culture and on the experiences of stakeholders is an essential aspect of critical evaluation. Through reflection and discourse, stakeholders can formulate explanations of the development of a program. By understanding the past, stakeholders will be in a better position to propose changes for program improvement.

Communication

As has been emphasized, discourse among stakeholders is a key feature of critical evaluation. These intense discussions among stakeholders and either external evaluators or designated internal evaluators involve self-reflection and are intended to reveal unexamined motives, meanings, and intentions.

Stakeholder Involvement

Critical evaluation relies on and promotes the democratic participation of stakeholders. Vocational education is a particularized program that reflects the interests, motivations, and needs of a complex and diverse group of stakeholders. Many of the benefits of extensive, in-depth involvement of stakeholders are long-range. Because of stakeholder involvement, fundamental changes in stakeholder attitudes can result in significant changes in programs.

Continuing Evaluation Effort

Critical evaluation does not have discrete beginning and ending points. It is a continuous process in which problem clarification, problem interpretation, and problem solving always interact. As the context changes, the problem changes, the information needed to examine the problem changes, and the most feasible solution may change. The selection of a solution may create conditions that will contribute to other problems. Program improvement is not a static affair. It demands attention through an interactive, continuing evaluation process.

Other distinguishing features of critical evaluation can be cited (e.g., capacity to use information collected and analyzed from the positivist and interpretative traditions, continuing educative function); those discussed here are the most important ones. A careful study of the features that distinguish critical evaluation from current evaluation efforts in vocational education reveals the enormous potential contribution critical evaluation can make to program improvement.

Constraints on the Conduct of Critical Evaluation

Probably because critical theory is only beginning to be taken seriously by American scholars, there are few examples of the actual conduct of critical evaluations (for some examples, refer to Sirotnik and Oakes [1983] and Lemish [1982]). Nevertheless, there are some factors that can be expected to act as constraints on attempts to conduct critical evaluation activities. One of the major of these is the extensive information requirement that most local vocational education programs are already having to meet. That is, local vocational education programs are already involved in meeting State and Federal evaluation and information requirements. In a real practical sense, these demands for accountability must be given high and even first priority. Moreover, as was suggested earlier, the interest in critical efforts to improve programs and change schools may very often be incompatible with the technical interest in managing and administering public educational programs. In such cases, critical evaluation activities may very well be considered "luxuries" that vocational education cannot afford. A legitimate question to be asked is whether or not it is reasonable or feasible to think that local vocational educators will take on yet another evaluation function that requires more information gathering and analysis. This raises another potential constraint on the conduct of critical evaluations--namely, the constraint posed by limited resources of time.

Clearly, critical evaluation activities require a substantial amount of time to implement and sustain. Thus, vocational educators considering the implementation of critical evaluation should recognize that time is a critical factor inherent in this approach. However, there is no magic answer to the question, How much time will be required? Time requirements depend on the complexity and diversity of the program, the nature of the problems studied, and so forth. Given that time is a scarce resource for teachers and school administrators, the key to the success of critical evaluation is administrative support. If the district and building principal do not make release time and rewards available for staff involvement, then the commitment needed for critical evaluation success may very well be insufficient. The time required should be viewed as a solid investment in improved programs.

Another constraint is related to the nature of democratic and participatory processes. In school settings where administrators are directive and coercive, critical evaluation may pose a threat to their position of control. For critical evaluation to achieve optimal results, it is essential that school administrators create an open and noncoercive environment, giving access to information and encouraging discussion of all aspects of evaluation problems. Stakeholders must be convinced that any comments they make will not be held against them. Participants in the evaluation process also need to understand how their recommendations will affect decision making and policymaking.

CHAPTER 7

SUGGESTIONS FROM EDUCATIONAL RESEARCH

This chapter provides a brief overview of the recent research literature in the areas of curriculum, school-based learning theory, effective schools, and organizational theory. The research conducted in these areas is rich in suggestions regarding how the schooling process may be improved, with significant implications for the theory and practice of vocational education evaluation. Since the delivery of vocational education is exercised through school teaching and learning situations, it seems reasonable to suppose that consideration of evaluation practice should draw on these areas for suggestions. Heretofore, this has not been the case. Present evaluations of vocational education tend to operate within a rather narrow perspective of the total schooling process. We hope in this chapter to demonstrate what educational research may have to offer vocational education evaluation and what it may imply for a critical evaluation framework.

Curriculum Theory: The Behavioral Objectives Approach

Curriculum has long been an object of scrutiny in education. However, it was only in the early 20th century that curriculum came into its own as a separate and distinct subject matter. As such, it has undergone the same amount of pluralism of approaches and opinions that other social scientific fields have experienced. Thus, in 1982, McCutcheon introduced a collection of curriculum papers in Theory into Practice by saying, "The authors whose work is included in this issue represent different positions in regard to what, how, and uses of curriculum theory, making for interesting comparisons; they also portray the field as one in confusion" (p. 2).

Walker (1983) explained the diversity of types of curriculum theory:

> We must acknowledge that there are important, fundamental differences among theories, even among the classics. When we talk of curriculum theory, we should use the plural. We must think of a family of theories with different purposes and forms bearing on the same problems. This diversity is not likely ever to vanish because each type of theory takes its own vantage point, each of these vantage points has validity and importance for some situations, and each appeals to some consumers. We must not be deluded into a search for a single type to which all theories should conform. (p. 64)

Based on a review of the traditional works in curriculum theory, Walker (1983) distinguished primarily four ways in which the phrase "curriculum theory" is used. The first of these is curriculum theory as "program rationalization." "This . . . type of curriculum theory proposes content, aims, and approaches to education--in short, it provides a program. It describes the program in detail and justifies it by giving reasons why it should be good and should be adopted" (p. 63). A second use of the term curriculum theory is in "procedure rationalization." Curriculum theory of this type "rationalizes procedures for curriculum construction or curriculum determination, rather than rationalizing the program itself" (p. 63). A third type of curriculum theory "conceptualizes curricular phenomena. This type is more removed from the immediate task of curriculum making. It sets out to advise those who directly address curriculum problems on helpful ways of thinking about the work. John Dewey's most influential writing on the curriculum takes this form" (p. 63). The fourth type of curriculum theory is that which "explains curricular phenomena . . . the fourth type frankly pursues understanding, leaving the application of the ideas to practice for others" (p. 64).

Walker's discussion of curriculum theory indicated the breadth of scope that the subject matter of "curriculum" encompasses. Moreover, each of these types of curriculum theory described by Walker has a significance of its own for evaluation and, thus, for program improvement. However, until recently, one particular theory or approach to curriculum design and construction dominated discussion in the field. This theory was the "behavioral objectives" approach formally introduced by Tyler (1950) in his now classic work Basic Principles of Curriculum and Instruction. Very simply, the behavioral objectives approach contains four broad areas of development activities. The first area is that of formulating the behavioral objectives the curriculum is intended to attain. The second is that of determining the content of instruction (i.e., the subject matter and information serving as the focus of instructional activities). The third area is the specification of the type and structure of learning activities that should occur. Last is the design of an evaluation of the curriculum's performance in meeting the specified objectives. Next, we discuss the areas of specifying objectives and evaluative methods since these have been the most controversial components of the behavioral objectives approach.

The behavioral objectives approach has its roots in the behaviorist movement. It will be remembered that a basic tenet of the behaviorist movement is that only those phenomena reducible to observational terms are appropriate objects for scientific inquiry. Unobservable phenomena or phenomena that can not be operationalized either do not exist in reality or belong to some other domain of inquiry (e.g., philosophy). In studies of individuals or groups, the phenomena most conducive to scientific investigation are human behaviors. Thus, theoretically speaking,

the specification of objectives in education "should describe both the kind of behavior expected and the content or the context to which that behavior applies" (Taba 1962, p. 200). Moreover, complex objectives such as providing students with understanding or providing students with competency must be broken down into more specific, separate, and discrete behaviors.

On the practical level, translating objectives into behavioral objectives ensures that objectives will be explicit and specific rather than vague and general. This is important in order to help teachers know what is expected of them. It is also important because the objectives can then function to guide "decisions about the selection of content and of learning experiences and of providing criteria on what to teach and how to teach it" (ibid., p. 197). Thus, objectives such as "teaching students to be democratic citizens" must be broken down into specifications regarding expected behaviors and the conditions under which such behaviors are expected to occur. Only in this way will it become clear how different subjects and different learning activities are expected to contribute to attainment of the objective.

In looking at the behavioral objectives approach, an immediate question comes to mind. On what basis do curricular developers select particular behavioral objectives? Where do they go to identify those objectives to be translated into desired and anticipated behavioral outcomes? Typically, four areas have been identified as appropriate sources of educational objectives. These include (1) assessments of students' needs, (2) various subject matters or disciplines, (3) educational philosophy, and (4) empirical studies of the culture within which the educational system functions.

There are several ways to talk about human needs. First, we can talk about inherent needs that individuals have as higher order living beings. Maslow's theory of psychological motivation is an example of the interpretation of such needs. Most often, need is usually defined as a gap that exists between some predetermined norm and the actual status of the individual. Examples of this type of need can be found in the renewed concern over and emphasis on basic academic skills; this is due to empirical evidence (i.e., test scores) that large numbers of students do not demonstrate an acceptable level of mastery in these areas.

Another major source of behavioral objectives is the various disciplines or subject matters. Indeed, Zais (1976) stated that subject-matter specialists are "probably the most commonly used source" (p. 304) of educational objectives. The disciplines may be related in at least two important ways. First, the various disciplines identify the understandings and skills necessary to acquire a specialized knowledge of the discipline's subject matter. For example, those in mathematics can inform curricula developers of the behavioral outcome students must attain for

math proficiency. Second, the disciplines have generated specialized knowledge that can be used to determine the behavioral outcomes necessary to attain more general educational goals. Thus, those in sociology may inform curricula developers regarding the specific behaviors that students must acquire to have healthy interpersonal relations in groups.

A special case of the disciplines serving as sources for educational ojbectives is that of philosophy and, more particularly, educational philosophy. Although the other disciplines may provide curricula developers with the concepts, methods, and findings specific to their subject matters, philosophy provides the foundations for particular value or ethical positions. The latter informs us as to what may be considered inherently good and, thus, desirable. Zais (1976) noted that "empirical sources have told us what is, but a philosophical inquiry is necessary to help us decide what ought to be" (p. 304). Educational philosophy determines both what life qualities are to be taken as inherently "good" and how the goals or aims of education can be formulated to contribute to that good. These determinations can then inform curricula specialists in the development of more particular and specific behavioral objectives to bring about the specified goals of education.

The fourth source encompasses the findings emerging from empirical studies of the dominant culture within which education functions. One major role often attributed to education is to provide students with the knowledge and skills necessary for conducting their lives outside the classroom. Thus, upon analyzing cultural content, it may be found that successful integration within the American culture requires individuals to have affective and cognitive skills conducive to tolerance for opposing viewpoints or, for example, skills needed to fit into the authority structures of typical work settings. Therefore, empirical analyses of the society's cultural content may provide the background for identifying specific behavioral outcomes needed for effective cultural integration.

As noted earlier, one of the major advantages attributed to the behavioral objectives approach is that it specifies clearly and distinctly what the outcomes of the curriculum are expected to be. The core evaluation question becomes, To what extent does the curriculum produce the desired behavioral outcomes? Since outcomes are usually operationalized in terms of learner behaviors, the key focus of most evaluation activity has been on student achievement in specified knowledge and skill areas. Most evaluation of behavioral objectives is product rather than process in nature. Also, most evaluations have been product rather than process in their orientation.

Zais (1976) has referred to this as the technical model of evaluation. According to Zais, this model specifies that "learners, viewed as 'raw material,' are subjected to certain

curricular and instructional treatments in order to produce a 'finished product' that meet predetermined objectives" (p. 370). Evaluation, then, consists of assessments of the curriculum's "productivity" in producing the finished products. Evaluative research may then be conducted on a comparative basis to determine the differential levels of efficiency achieved by different curricula. Assessment of the curriculum's productivity is generally conducted through measurement in quantifiable terms of changes in students' behaviors. The most common form of measurement undertaken is that of student testing. Data received as a result of student testing can then be measured against some standard to determine the curriculum's degree of effectiveness (i.e., productive) in producing the desired changes in students' behaviors.

Although the behavioral objectives approach may be losing ground in general education, variant forms of it still predominate in vocational education. One form is actually taken from the forerunner of the Tyler rationale. This is the job and task analysis approach adapted from Bobbitt's application of time-and-motion studies to the curriculum development area. The second form is the competency-based or performance-based education approach. These are briefly discussed in the following sections.

Job and Task Analysis in Curriculum Construction

Kliebard (1977) refered to the job-task analysis approach as a "production" model of curriculum development. MacDonald (1975) called it the "technical" model and claimed that the language of job-task analysis is "fundamentally grounded in a technological rationale rather than in philosophical and/or scientific theory" (p. 7). He then stated that "this technical model has been developed to its greatest sophistication by vocational education workers" (p. 7).

Job analysis, according to Larsen (1969), is the "collection and interpretation of information about the work performed. . . . [It] is needed to describe the job toward which training is directed and to determine the effectiveness of training as reflected in job performance" (p. 21). Although different methods of conducting job analyses are available, they all share the common aim of developing an "inventory" of the learning activities associated with specific instructional areas for specific jobs.

Once a job analysis has been conducted, a task analysis determines the knowledge and skill requirements of the job, as indicated by Larsen (1969):

> Basic to task analysis is the identification of the kinds of performance capabilities demanded by the tasks. Each task must be analyzed to determine the basis for all decisions. The selection of appropriate

> objectives, content, sequence, method, media, and evaluative criteria depends on the correct identification of the capabilities needed to perform the tasks. (p. 23)

Smith and Mass (1970) provided an overview of the process of curriculum development based on job-task analysis. Summarized, the steps in the process include (1) specification of the job role for which training is to be provided, (2) identification of the tasks conducted in the specified role, (3) selection of the tasks to be taught, (4) analysis of each of the selected tasks, (5) a statement of the performance objectives students will be expected to attain, and (6) specification of the instructional program and sequence to be implemented to achieve the required level of job performance.

Central to the notion of job-task analysis is the application of scientific (i.e., positivist) methodology in the specification and analysis of job roles and associated tasks. Thus, empirical methods are used (e.g., surveys, controlled observation) to determine specific job roles and the particular tasks involved in fulfilling these roles. Moreover, it is anticipated that a thorough job-task analysis can provide the basis for a curriculum that is generalizable to diverse contexts and employment situations in which similar job roles are required.

Competency-based Education

Closely related to job-task analysis is the competency or performance-based approach to curriculum development. This approach's close association with job-task analysis and the Tyler rationale is illustrated in the writings of Hall and Jones (1976).

> Competency-based education, first of all, is education that focuses on students' acquisition of specific competencies. In other words, the educational program includes a set of learning objectives that are stated so that their accomplishment can be observed in the form of specified learner behaviors or knowledge. Minimum levels of achievement of these objectives are established as a criterion of success. Learning activities are geared to assist each student in acquiring at least the minimum levels of competency. (p. 11)

Norton (1984) listed what he considered to be the five essential (and desirable) characteristics of competency-based education.

- Competencies to be achieved are carefully identified, verified, and made public in advance.

- Criteria to be used in assessing achievement and the conditions under which achievement will be assessed are explicitly stated and made public in advance.

- The instructional program provides for the individual development and evaluation of each of the competencies specified.

- Assessment of competency takes the students' knowledge and attitudes into account but requires actual performance of the competency as the primary source of evidence.

- Students progress through the instructional program at their own rate by demonstrating the attainment of specified competencies.

Norton seems to be describing a particular kind of competency-based education. His attention to instructional strategies limits the generalizability of his description. Although individualized instruction may be a desirable instructional strategy for implementing competency-based curricula, it is not the only acceptable strategy.

Competency-based curriculum design has much in common with the behavioral objectives and job-task analysis approaches, with one very important difference. This difference lies in the fact that, although behavioral objectives and job-task analysis curricula are generally geared for groups of students, competency-based education is geared to the individual student. Thus, as Norton (1984) pointed out, in competency-based curricula, "learning is the constant and time the variable" (p. 1) rather than the usual case where instructional time is the same for a group of students resulting in variable achievement levels for individual students. Competency-based curricula allow the selection of instructional strategies that permit the individual student to take whatever time is needed in order to attain the prescribed criterion level of achievment. Thus, ideally, no students are discriminated against due to earlier learning inadequacies or current learning problems.

Critique of the Behavioral Objectives Approach

A number of strengths or advantages have been attributed to the behavioral objectives approach, including job-task analysis and competency-based education. Briefly summarized, some of the major strengths are described as follows:

- All three approaches have a "built-in" component or basis for evaluating the curriculum, which is the product of the process. This is because all three

approaches include the specification of "objective" standards by way of which the quality of the curriculum and instructional activities can be assessed as effective or ineffective. These objective standards are created in the explication and specification of the behavioral competencies and outcomes that students are expected to demonstrate after receiving the particular educational treatment.

- Moreover, each of these three approaches has "rationalized" the curriculum construction process by composing it into discrete and functional components (i.e., the specification of behavioral objectives, job roles, or desired competencies, the specification of appropriate instructional sequence and activities, and so on). Classic management theory suggests that rationalization is closely related to quality control and usually renders human-designed processes both more effective and more efficient.

- A third advantage of each of these processes is that their end products are generalizable with little or no adaptation to a large number of diverse contexts and learning situations. Thus, they are cost-effective insofar as use of their products prevents educators from having to develop or to revise curricula for individual students or individual classes. The generalizability of the curricula that result from these processes is reflected in the creation and marketing of numerous types of learning packages and modules for varying educational and training objectives within the last decade.

- In terms of vocational education per se, the job-task analysis approach is seen as particularly advantageous since it has a high probability of ensuring that students are, in fact, being trained in those skills needed by business and industry. This is clearly because the identification and specification of job roles and corresponding tasks and competencies are based on an empirical study of the tasks employees currently perform in business and industry.

- Finally, competency-based curricula have the advantage implied earlier of allowing individuals to proceed through a step-ordered learning sequence at their own pace. This is advantageous not only because it enhances individuals' opportunities to learn and acquire skills but also because it is cost-effective in expenditure of educational resources. Its cost-effectiveness lies in the fact

> that the units of time spent on individual learning activities are more directly tied to attainment of individual learning outcomes. This is in contrast to group-paced curricula in which group learning time may be effective for only a small percentage of the class.

Although numerous advantages of the behavioral objectives approach, including job-task analysis and competency-based education, have been delineated by proponents for such a curriculum, criticisms of this approach are beginning to mount. Zais (1976) noted that behavioral objectives are not applicable "when it comes to higher-order tasks (e.g., analytic or synthetic thinking, internalizing of value systems, or appreciation of art and literature)" (p. 313). His argument for this claim is based on the reality that it is not practically possible to list all the observable behaviors for these complex constructs. Even if this list were constructed, Zais (1976) argued that such a practice would treat education as a closed-end enterprise.

> The human organism, behavioral critics claim, is not a finished product, and to define him in terms of prescribed behaviors is to ignore the most signifgicant question that the curriculum poses: What can man become? (p. 314)

Furthermore, one must ask, Who defines the objectives or competencies? Often they are defined at the top by curricula experts and are given to teachers for implementation with little or no allowances for individual school-setting adaptation or modification. In essence, the curriculum becomes an instrument for technical activity that excludes the participation of those for which it is intended. Moreover, teachers become an instrument of the curriculum instead of the converse. If teachers are supposed to be central to the educational process, it seems that behavioral objective curriculum negates their ability to utilize their practitioners' knowledge, talents, and experience. In effect, the curriculum dominates students and teachers instead of actualizing their human potentials.

Inherent in this curriculum discussion is the training-education dichotomy that exists in all school situations. According to Zais (1976), training "is referred to as the technical model in curriculum construction, while (liberal) education sometimes is called the humanistic model" (p. 316). Although curriculum seldom is classified as either purely training or education, particular curriculum may emphasize one or the other. In fact, vocational education is often referred to as vocational training due to its highly structured, behavioral outcome approach designed for job competence. Although vocational training is an integral part of vocational education, this manuscript suggests that training should not dominate the curriculum in the

public school setting. In summary, vocational education curriculum should be more than an instrument for delivering job competencies.

Although numerous criticisms of the behavioral approach to curriculum have been leveled, no alternatives have been universally accepted. However, there is a growing body of literature that is responding to these criticisms. Not surprisingly, the individuals who are suggesting alternatives come from orientations that include phenomenology, educational philosophy, political science, and critical theory. Scholars such as Aoki and Harrison (1979), Bowers (1984), Cherryholmes (1983), and Feinberg (1983) envision curriculum as a medium that has normative and practical content as well as technical content. Moreover, these scholars suggested that the structure of the meaning of schooling is taken for granted instead of being critiqued through critical discourse.

In summary, given the limiting effect that the technical model of curriculum has on education, it appears that vocational education also needs to question its structure of meaning in light of the total schooling process. Although it is important that students acquire technical competence, the importance of normative and practical content cannot be slighted. Curriculum that separates vocational education and general education from the total schooling process ignores the gestalt of the total schooling experience. Critical discourse is a step in the right direction for reconceptualizing vocational education curriculum so that present curriculum deficiencies can be ameliorated.

School-based Learning Theory

In 1963, Carroll (1963) introduced his now classic "model of school learning." Central to the model was the notion that learning takes time and that different learners require different amounts of time to master new learning tasks. More specifically, in the writing about the Carroll model, Bloom (1982) indicated that

> if students are normally distributed with respect to aptitude for some subject and all students are given exactly the same instruction (the same in terms of amount and quality of instruction and learning time allowed), then achievement measured at the completion of the subject will be normally distributed. Under such conditions the relationship (correlation) between aptitude measured at the beginning of the instruction and achievement measured at the end of the instruction will be relatively high. . . . Conversely, if students are normally distributed with respect to aptitude, but the kind and quality of instruction and learning time allowed are made appropriate to the characteristics and

needs of each learner, the majority of students will achieve mastery of the subject. And, the correlation between aptitude measured at the beginning of instruction and achievement measured at the end of instruction should approach zero. (p. 4)

The most noted extension of the Carroll model has been the school-learning model proposed by Bloom (1982) in his work Human Characteristics and School Learning. Although a full explication of Bloom's model lies beyond the scope of this work, a review of some key points regarding the model will indicate its implications for educational and vocational education evaluation.

Bloom's School-based Learning Model

In his first chapter, Bloom (1982) emphasized: "The main thesis of this book is that individual differences in learning is an observable phenomenon which can be predicted, explained, and altered in a great variety of ways. In contrast, individual differences in learners is a more esoteric notion" (p. 8). Bloom assumed that differences in achievement among learners are a product of variables in the learner's environmental context and not necessarily the product of genetic traits that are fixed in the individual at conception.

The second major assumption underlying the Bloom model is that

> one of the most important elements in accounting for individual differences in school learning is the centrality of instruction for groups of learners. Instruction provided to a group of twenty to seventy learners is likely to be very effective for some learners and relatively ineffective for other learners. This aspect of the process of schooling is likely to be replete with errors which are compounded over time. (Ibid., p. 9)

Thus, the model assumes a small number of variables in the schooling process can be identified that account for a great deal of the variation in school learning. The identification of these variables may, in turn, contribute to the design of "minimal-error systems" of schooling.

The model shown in figure 2 includes three interdependent variables (or classes thereof).

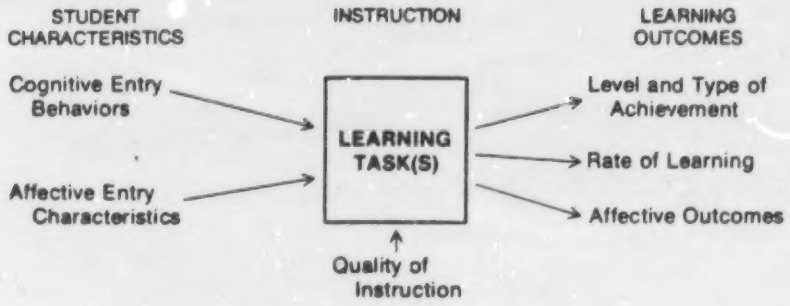

Figure 2. Major variables in the theory of school learning

SOURCE: Bloom (1982, p. 11)

Bloom (1982) defines a learning task (the basic unit of analysis) as a "learning unit in a course, a chapter in a text book, or a topic in a course or curriculum. . . . This type of unit contains a variety of ideas, procedures, or behaviors to be learned over a relatively short period of time" (p. 22).

Under "student characteristics," cognitive entry behaviors refer to the prerequisite learning necessary for mastering a new task. Affective entry characteristics refer primarily to the learner's level of motivation to learn new tasks. Under "instruction," quality of instruction according to Bloom (1982) refers to "the extent to which the cues, practice and reinforcement of the learning are appropriate to the needs of the learner" (p. 11). Variations in and interactions between student characteristics and instructional variables cause different levels of outcomes among learners.

Research Findings on School-based Learning

Beyond explicating the model of school learning, Bloom's 1982 work also provides a review and synthesis of educational research pertinent to the model's major components. The reader is referred to Bloom's work for an extensive and detailed review of research findings. Next, we will briefly summarize some of the key findings related to the model's major variables.

Student characteristics. Regarding student cognitive entry behaviors, Bloom concluded that the evidence from both macro- and microstudies indicated that there is a strong correlation between

student cognitive entry behaviors and achievement on subsequent learning tasks. Bloom (1982) concluded that "cognitive entry behaviors are causal links in determining learning and in accounting for cognitive educational achievement" (p. 68). This conclusion is applicable to both specific cognitive entry behaviors and to general ones such as verbal ability, reading comprehension, and so on. Moreover, the studies indicated that when helped with cognitive entry behaviors, students perform better on subsequent learning tasks. It is less clear as to the impact that help with cognitive entry behaviors in later instructional sequences may have on achievement.

Review of the research on affective entry characteristics indicated that there is a significant relation between affective characteristics and related measures of learning achievement. The studies reviewed tended to focus on three types of affective characteristics. The first is the most narrow and is referred to as "subject-related affect." Subject-related affect is identified with how a student feels toward or likes a particular school subject (e.g., geometry, German, biology). "School-related affect" is a characteristic related to how students in general feel about or view schools and schooling. Finally, "academic self-concept" refers to students' perceptions of themselves in relation to the school-learning process (i.e., do they perceive themselves as successful learners or failures; do they perceive themselves as capable of learning new tasks or not). Studies indicate that these three affective characteristics appear to be separate and distinct in their impact on achievement in a learner's early school years. Moreover, they also appear to be amenable to change. However, they tend to become more generalized into a single dimension in later grades and also tend to become increasingly structured and resistent to change.

The studies also indicate that there is an interactive relation between affective entry characteristics and school achievement. Thus, whereas a student's affective dispositions will influence his or her achievements, the achievement experience itself will then affect the way in which he or she perceives various subject matter, the learning process, and him or herself as an efficacious learner. On the basis of this conclusion, Bloom (1982) suggested that, "teaching, curriculum, and grading policies in the school which stress high ratios of success experience to failure experiences should result in increased amounts of positive affective entry characteristics for subsequent related learning tasks." (p. 105)

Quality of instruction. At the outset of his discussion regarding the quality of instruction, Bloom (1982) noted that current research shows student characteristics having "far more affect" on learning than does quality of instruction. Indeed, evidence suggests that the quality of instruction cannot overcome the effect of negative cognitive and affective entry

characteristics unless (1) the instruction is aimed directly at changing those entry characteristics, or (2) the learning task is redefined so as to make it "learnable" by students with whatever entry characteristics they bring to the situation. Nevertheless, quality of instruction still cannot overcome inadequate entry characteristics for a particular learning task. (Bloom [1982] does suggest, however, that quality of instruction is more successful in overcoming negative affective entry characteristics than it is in overcoming negative cognitive entry behaviors.) Having drawn this conclusion, Bloom (1982) continued by saying that he thought quality of instruction may nevertheless be a powerful determinant of learning for particular learning tasks.

Bloom noted that much of the recent research has focused on the relation between school earning and characteristics of teachers, classrooms, and schools. Thus, a number of studies have been conducted that attempt to relate teacher age, salary, previous experience, teaching attitudes, and so on to learning outcomes. Other studies have looked at how size of class, availability of diverse learning resources, expenditure per student, and management policies affect student achievement. Review of this research led Bloom (1982) to draw this conclusion:

> There is at present little support for believing that the characteristics of teachers, classrooms, or schools have much effect on the learning of students. . . . We believe that it is the teaching and not the teacher that is central, and it is the environment for learning in the classroom rather than the physical characteristics of the class and classroom that is important for school learning. (p. 111)

Bloom then suggested that there has been an overemphasis on how teachers manage the learners, with little or no attention paid to how teachers manage the learning. According to Bloom, Carroll (1963) focused on the latter when he defined quality of instruction as "the degree to which the presentation, explanation, and ordering of elements of the task to be learned approach the optimum for a given learner" (p. 111). Bloom (1982) pointed out that the implicit assumption in this definition is that "each student can learn if the instruction approaches the optimum for him or her" (p. 111). Thus, although Bloom differs from Carroll with regard to what is involved in quality of instruction, they both share the assumption that students differ with regard to the quality of instruction they need to master a particular learning task.

What does Bloom think is involved in quality of instruction? Bloom (1982) suggested that one method of identifying the ingredients of quality instruction is by analyzing what occurs in the optimal learning situation of a tutor and a single learner.

> Stripped to its essentials, Quality of Instruction . . . has to do with the cues or directions provided to the learner, the participation of the learner in learning activity (covert or overt), and the reinforcement which the learner secures in some relation to the learning. Because much of school instruction is group instruction and because any attempt at group instruction is fraught with error and difficulty, a feedback and corrective system must be also included in the quality of instruction. (p. 115)

Briefly defined, cues refer to any communication that instructs learners as to what is to be learned as well as to direct students as to what they will do in the learning process. Reinforcements are generally any action or event that increases the probability that learners will provide appropriate responses to particular cues. Thus, reinforcements may be positive (e.g., rewards) or negative (e.g., punishments) in nature.

Participation refers to any form of active involvement of the learner in the learning process. Participation may mean rote practice, verbal responses to questions, silent studying of texts, and so on. In short, participation exists when the student is actively engaged in learning. Finally, feedback and correctives are essential components of the learning process. Their use enables the educator to determine the adequacy of cues, reinforcements, and participation for each student as well as the ways in which these instructional elements need to be revised. Bloom (1982) noted that "feedback procedures typically consist of brief formative tests, at the end of each learning task, which indicate what the student has learned, and what he still needs to learn to attain mastery of the task" (p. 125). Correctives, then, are the recommendations regarding what changes or alterations in cues, reinforcements, and participation should occur in order to bring a student up to mastery level. The primary emphasis in the corrective procedures is on "alternative cues and additional time and practice" (p. 125). Use of any and all of these elements in the learning process is what Bloom refers to as "mastery learning strategies."

In Bloom's (1982) review of research on the quality of learning, he found that the use of cues, reinforcement, and participation accounts for 20 percent of the learning variation in the macrostudies or correlational studies and 25 percent of the learning variation in the micro- or quasi-experimental studies. Moreover, in the macrostudies, participation stood out as the strongest single factor relating to learning differences. In microstudies conducted at the University of Chicago, the use of feedback and corrective procedures appeared to be especially effective in accounting for the learning differences of students randomly placed in mastery and nonmastery classes. It should be noted that these findings are reinforced by the "effective schools" literature in which level of participation and use of

feedback and corrective procedures have been cited as characteristics related to student achievement levels.

Another set of interesting conclusions noted by Bloom (1982) relates to the effects of quality of instruction over time.

> In general, we find that students who have developed effective leaning and study procedures will be less affected by variations in the quality of instruction than students who have developed less effective learning procedures. Similarly, more mature and capable learners will be least influenced by varying qualities of instruction. . . . Thus, the generalization appears to be that students can acquire learning procedures which will enable them to learn well under less than ideal qualities of instruction. (p. 135)

In addition, Bloom (1982) suggested that quality of instruction is a central influence on the history of the learner--history referring to the set of characteristics that a student brings to a particular learning task. In this sense, then, quality of instruction is also an important determinant of the individual's future learning capacity or ability. The implication of these findings and hypotheses is that mastery instructional practices can, over time, enhance students' learning capacity as well as their learning status at any one point in time.

Affective outcomes of school learning. Because the long-term ramifications for the individual and society are so significant, the subject of affective outcomes of school learning is given separate treatment. Bloom (1982) hypothesized that affective outcomes are "influenced directly by school achievement and especially by the student's perceptions about his school achievement" (p. 140). Moreover, while research suggested that achievement and affect are interactive, it also suggested that overtime achievement has prior and primary influence on affect.

In summary, Bloom theorized that in moving from one learning task to another, from one subject to another in a sequence, and from one school year to the next, students develop increasingly consistent and structured perceptions regarding their competency to learn. Such perceptions are significantly related (and become more so over time) to the public judgments or assessments (i.e., test and summative grades), regarding their relative status in learning performance. The research reviewed by Bloom tends to confirm this hypothesis in each of the separate areas of subject-related affect, school-related affect, and academic self-concept.

Finally, research also suggests that in schools or classes in which mastery learning processes are used, including additional time and help provided to individual students who need them, affective outcomes are generally more positive than in schools and classes in which less-favorable learning conditions

exist. These are significant findings because they suggest that schools and educators at the local level can influence over time not only learners' perceptions of themselves as learners but also their very capacity to learn. This in turn is of great importance because other studies indicate that the capacity and motivation to learn is a necessary component of the individual's ability to cope with change in modern and technological society. Moreover, other studies indicate that the capacity to learn and the individual's learning self-concepts are related to the status of their mental health. Thus, affective outcomes of the schooling process have a significance for people and society that goes beyond the matter of what is learned.

Synthesis of the components of the school-learning model and their interrelationships. Having reviewed research pertinent to each of the three major independent variables in the school learning model, Bloom (1982) attempted to analyze evidence regarding the composite model. In doing so, he primarily relied on findings obtained in the microstudies conducted at the University of Chicago.

> Students were placed into a mastery learning section and a control section of the course so that the two groups of students were roughly equivalent on measures of previous achievement and aptitude. Where possible, the same teacher taught both sections. The subject matter was the same, the learning materials used were identical, and the instructor(s) attempted to use similar instructional procedures. (p. 175)

In the control group, instruction on each of the learning tasks in the sequence was provided followed by a criterion-referenced test on the material covered in the task. However, test results were not reported to control group students nor did they receive any correctives (e.g., additional time or help) for inadequate performance. The mastery group also received instruction on the same tasks, followed by a test. However, in contrast to control students, students in the mastery group received communication regarding their test performance. Moreover, mastery students were provided with whatever time and additional instructional help needed in order to master the criterion in those areas missed on the test. Both the control group and the mastery group were given a summative test at the end of the learning sequence.

In general, analysis of the results of the microstudies support Bloom's school-learning theory. Thus, according to Bloom (1982), the findings indicated the following:

- The mastery group showed increasing achievement on cognitive measures throughout the learning series, whereas the control group actually declined in achievement. Moreover, as expected, the mastery

group showed increasingly less variation in the achievement over the course of the series than did the control group.

- In terms of affect, the two groups that were approximately similar in initial affective characteristics tended to be increasingly divergent (where we have measures at each learning task) at the last measure of affect.

- In terms of affect, the master group became increasingly differentiated from the control group in the direction of more positive affective characteristics. However, unlike the scores on cognitive achievement, the mastery group displayed equal or more variation in affect in two of the four studies in which affect was measured.

- Data on the groups regarding quality of instruction suggested that "less than perfect" matching of the two groups had occurred. Nevertheless, it still appeared that the mastery group learned increasingly more in each learning task than did the control group. Moreover, it was found that on summative measures for the learning series, the mastery group displayed only about 20 percent of the variation in achievement than did the control group.

- The combined effects of cognitive entry behaviors, affective entry characteristics, and quality of instruction accounted for less variance in achievement in the total group of mastery and nonmastery students than had been predicted. However, this still came to a respectable correlation of + .70.

- In reviewing these findings, Bloom concluded: "What these figures indicate is that the effect of positive entry characteristics (cognitive plus affective) and improved quality of instruction is to increase the relative learning effectiveness of the mastery students, while the less positive entry characteristics and less favorable quality of instruction decrease the relative learning effectiveness of the control students when they encounter successive new learning tasks in a sequential series" (p. 186).

Bloom (1982) also reported the findings of 12 macrostudies that analyzed differences between mastery and nonmastery classes in their natural school environments. Although most of these studies reported that the mastery classes had significantly

higher means of achievement than did nonmastery classes, a number of them also reported that the median variation of the mastery class in contrast to the control class was about 60 percent. This is considerably higher than the 20 percent reported for the microstudies. These findings led Bloom (1982) to this conclusion:

> Evidently, under regular school conditions, mastery learning can produce sizeable achievement differences even though it does not provide the most favorable conditions for learning and reduction of individual differences that we believe possible. (p. 198)

Bloom's model is a noteworthy attempt to capture the complexities inherent in the school-learning process. The notion that almost all persons can learn if they are given appropriate prior and current conditions of learning seems logical. Moreover, Bloom's contention that greater degrees of learning equality can be achieved and that students can approach equality of achievement while, concommitantly, approaching equality of time needed to learn has an intuitive appeal. In essence, Bloom's model proposes an egalitarian dream: equality of opportunity (time) and of outcomes (achievement) at levels of excellence.

Although there are numerous variables that comprise Bloom's school-learning theory, interestingly, time on task as it relates to student achievement recently has received increased attention. The literature as it pertains to academic classes suggests that the amount of time students are engaged in learning relates positively to their academic achievement. As a result of this recent attention, Halasz and Behm (1983) conducted a time on task study in the vocational education school setting to determine the proportion of time spent on selected content areas. Although they assumed that time on task is positively correlated with learning achievement, Halasz and Behm reported that "there is no formula for calculating the precise amount of time required for optimal learning at either the elementary or secondary level" (p. 9). The point where more learning time does not produce more learning, or negates prior learning, is different for each student. Although Bloom's model in total would acknowledge this threshold, the recent emphasis on time on task seems to single out the time factor for investigation while ignoring its relationship with other variables in the model.

It appears that equality of learning time and equality of achievement are mutually exclusive. If equality of achievement is chosen as a goal for vocational education, then inequality of time is needed to obtain that goal. Consequently, slower students would require more time to master program units and would eventually find themselves falling behind those who learn at a faster rate.

Critique of Bloom's School-based Learning Model

Although Bloom's model of school learning has not received criticisms of any significant amount, it is clearly a technological-theoretical model similar in nature to the technological frameworks criticized earlier in this manuscript. According to Bloom's model, the learning process is fragemented into discrete variables. Learning outcomes are dependent upon antecedent variables. In essence, the model ignores the student as a person by focusing attention on the input-output process that drives the model. Somewhere in this process, the student as a person is transformed into a behavioral outcome that is quantified. Moreover, contrary to Bloom's belief that the model is value-laden, it does imply prescriptive procedures for learning which, in turn, reflect a value position. In summary, given the technological framework of the model, it is assumed that critics of the technological frameworks discussed earlier in this manuscript would be critical of Bloom's model as well.

In fairness to Bloom's model, technical knowledge as a principle per se is not bad. However, one needs to ask how this model can be extended to incorporate the student as an active actor in the learning process instead of reducing the student to an object of repetition and practice. In other words, the model should encourage student participation so that they would be treated as self-defining subjects in the learning process. In conclusion, it appears that Bloom's model is rich in suggestions for further evaluative inquiry and critical discussion. The interrelation of variables that Bloom proposed provides a useful framework for initiating critical discussions on the complex schooling process. The challenge is to refashion this model, or a version of it, to make it more amenable to critical discourse and normative and practical applications.

Effective Schools Research

The school achievement research literature is replete with studies that are pessimistic about the effectiveness of public schools. The Coleman report entitled Equality of Educational Opportunity (1966), Jensen (1969), and Bowles and Gintis (1976) exemplify studies that contain negative findings on school effectiveness. In general, these studies concluded that schools make little difference in achievement as compared to family environment.

The pessimistic tone of these findings did not go unchallenged. The Rand Corporation (Averch et al. 1972) identified four substantive problems encountered in virtually every area of educational research:

- The data used by researchers are, at best, crude measures of what is really happening. Researchers in education are plagued by the virtual impossibility of measuring those aspects of education they wish to study.

- Educational outcomes are almost exclusively measured by cognitive achievement, despite the many serious problems involved in interpreting such scores.

- There is virtually no examination of the cost implications of research results. This makes it difficult to translate research results into policy-relevant statements.

- Few studies maintain adequate control over what actually goes on in the classroom as it relates to achievement. (p. 18)

MacKenzie (1983) noted:

> All along, some researchers had been looking at schooling as an open-textured process of instruction and learning rather than a closed system for the distribution of limited outcomes (Jackson 1968, Leacock 1969). Thanks in part to the spirited debates over negative findings in evaluation research (Cicirelli, Evans, and Schiller 1970; Hellmuth 1970; House et al. 1978; Smith and Bissell 1970), new studies started to look much more carefully at the processes embodied in educational interventions as opposed to the labels and formal prescriptions attached to them (Cooley and Leinhardt 1980) (p. 6)

Cooley and Leinhardt (1980), in their instructional dimensions study sponsored by the National Institute of Education, focused on effective classroom processes in regular classroom settings. Their study is an example of an approach to program evaluation that considered educational treatments a multidimensional domain instead of using the traditional approach of employing simple contrasts in quasi-experimental design.

The advent of Carroll's (1963) process model of school learning that was eventually furthered by Bloom (1976) added impetus for the need of process approaches to educational research and program evaluation. As noted by MacKenzie (1983):

> Process models provide a flexible theoretical base for program evaluation (Cooley and Leinhardt 1980) and for the study of teaching (Denham and Lieberman 1980).

> Constructs such as learning time, readiness, and likelihood of success help to specify the "alterable variables" within the school environment (Bloom 1976). (p. 6)

The focus of recent effective school studies examine the school in the total setting. MacKenzie in his 1983 review of the recent effective schooling literature, perhaps best summarized the research results:

> All the new studies agree that schooling is a complex and continuous, multi-faceted process that is always conditioned by the history and circumstances of its evolution. As a consequence, no single element of school effectiveness can be considered in isolation from all the others, or from the total situation in which it is found. The principles of effectiveness may be consistent, but each school must implement them in unique ways. When effective schools are examined in vivo, what emerges is not a checklist of specific ingredients but a "syndrome" or "culture" of mutually reinforcing expectations and activities (Purkey and Smith 1982). (p. 8)

Furthermore, from his review of the effective schools' literature, MacKenzie identified 31 elements that aid in the understanding of school effectiveness. He clustered these elements under three dimensions: leadership, efficacy, and efficiency. Each dimension consists of core and facilitating elements. Looking at figure 3, MacKenzie (1983) noted:

> As we move down the list from leadership to efficacy to efficiency, we also move progressively from the school and district level of implementation toward the individual classroom. Ultimately, here is where the learning takes place by which we will judge a school's effectiveness. (pp. 8-9)

Later, MacKenzie stated:

> What I have clustered as efficiency dimensions all have to do with how time spent in the classroom is used to promote learning. (p. 9)

Goodlad (1984), in A Place Called School claimed that the continued practice of tracking will ensure ineffective schools. Intuitively, tracking appears to make sense. The grouping of students by ability or achievement so that individual differences can be better accommodated sounds logical. However, Goodlad (1984) attacked this popular assumption of tracking: "In effect, however, it serves as an organizational device for hiding awareness of the problem rather than an educative means for correcting

LEADERSHIP DIMENSIONS:

Core Elements
Positive climate and overall atmosphere
Goal-focused activities toward clear, attainable and relevant objectives
Teacher-directed classroom management and decision making
Inservice staff training for effective teaching

Facilitating Elements
Shared consensus on values and goals
Long-range planning and coordination
Stability and continuity of key staff
District-level support for school improvement

EFFICACY DIMENSIONS:

Core Elements
High and positive achievement expectations with a constant press for excellence
Visible rewards for academic excellence and growth
Cooperative activity and group interaction in the classroom
Total staff involvement with school improvement
Autonomy and flexibility to implement adaptive practices
Appropriate levels of difficulty for learning tasks
Teacher empathy, rapport, and personal interaction with students

Facilitating Elements
Emphasis on homework and study
Positive accountability; acceptance of responsibility for learning outcomes
Strategies to avoid nonpromotion of students
Deemphasis of strict ability grouping; interaction with more accomplished peers

EFFICIENCY DIMENSIONS:

Core Elements
Effective use of instructional time; amount and intensity of engagement in school learning
Orderly and disciplined school and classroom environments
Continuous diagnosis, evaluation, and feedback
Well-structured classroom activities
Instruction guided by content coverage
Schoolwide emphasis on basic and higher order skills

Facilitating Elements
Opportunities for individualized work
Number and variety of opportunities to learn

Figure 3. Dimensions of effective schooling

SOURCE: MacKenzie (1983, p. 8).

it. The decision to track is essentially one of giving up on the problems of human variability in learning. It is a retreat rather than a strategy" (p. 297). Although the research literature on the negative side effects of tracking abounds, these findings have had a difficult time receiving a hearing. Goodlad (1984) concluded that tracking needs to be eliminated. He believed that nearly all students can learn if given the appropriate support, corrective feedback, and time.

Furthermore, Goodlad (1984) was critical of the vocational-academic split. He opposed the common belief that a significant part of the population was destined to have serious difficulties in school and, thus, should be partitioned into vocational education. On the contrary, Goodlad advocated vocational education for all students instead of its being "just an alternative to academic studies for the less academically oriented" (p. 147).

> I want college-bound students to include vocational education studies too, just as I want to be sure that students not going to college secure a balanced program in academic subjects. So long as vocational education is an elective of doubtful status and difficult accessibility for students preparing for college and not one of the five fingers of the hand described by the Harvard Report--mathmematics and science, literature and language, society and social studies, the arts and vocations--it will sit somewhat apart from the rest of the curriculum and have less than first class status. (p. 148)

Moreover, Goodlad (1984) noted, "There are evidences in our data, growing stronger in the senior highs, of students' and parents' awareness of careers and the role of school in them. For both groups, the importance of vocational education increases" (p. 144).

Evaluators of vocational education programs need to heed the recent research findings on effective schools. Suggestions for vocational education direction and for vocational education program evaluation begin to emerge from these findings. The findings suggest that if vocational education is to contribute to making schools more effective, vocational education program evaluations must encompass the complexity of the teaching and learning environment. Reliance on outcome measures alone ignores the processes that effect those outcomes.

Organization Theory

The current research from organization theory and design, organizational assessment, human resource development, and program implementation is supportive of the need for a process

such as critical inquiry to facilitate the process of change aimed at organizational effectiveness. A review of research findings outside the school area, primarily in the business and human services area, is appropriate as these areas are focusing more and more on qualitative methods of organizational assessment than is vocational education.

A number of controversies have surfaced over the years as to the importance of various elements for organizational effectiveness. With the broad acceptance of the general systems approach of organizational functioning, many of these issues have come to be regarded as part of an integrated and multifaceted social system. Research orientations have also shifted from macromechanistic models of "rational" man to more organic models of individual, group, and organization-specific interconnectedness. Of particular importance to the notion of critical inquiry is the movement away from the "closed system" concept of organizations to the "open system" concept where the ability of the organization to interact with and adapt to the external environment is a critical component for its survival.

A summary of findings from these areas along with the relevance to education is presented next.

Organization Theory and Design

Daft (1983) listed 11 areas that have been studied in varying degrees and that have some relevance to organizational theory and design. He divided them into areas of high, moderate, and low development depending on the level of understanding and agreement about the specific framework.

TABLE 1

AREAS OF HIGH, MODERATE, AND LOW DEVELOPMENT IN ORGANIZATION THEORY

HIGH	MODERATE	LOW
Bureaucracy & size	Environment	Power and politics
Technology & structure	Goals & effectiveness	Information and control
	Innovation & change	
Functional vs. product structure	Decision making	Top management domain
	Intergroup relations	

SOURCE: Daft (1983, p. 507).

Bureaucracy and Size

There is a positive relationship among size, bureaucracy, and performance, where bureaucracies are better suited to stable and moderately changing environments. Performance is not always improved by increasing bureaucratic characteristics. This strategy should not be used when organizations are small or when employees are highly professional, or when the environment is unstable. This aspect is of particular importance to educational institutions in which the administrative core is generally bureaucratic and the "production core," composed of professionals, is more autonomous. This is similar to Weick's (1976) concept of educational institutions as "loosely coupled systems."

Technology and Structure

Technology has a relationship with specialization, standardization, formalization, and centralization. Although size is a better predictor of most structural characteristics, technology has a stronger relationship in smaller organization. Woodward (1965), Thompson (1967), and Perrows (1970) developed various typologies dealing with various aspects of technology. For example, Perrows's framework was applied to department technologies within organizations. Understanding the variety and analyzability of a technology tells which management style, structure, and process should characterize that department to enhance satisfaction, effectiveness, and efficiency.

Functional vs. Product Structure

Organizational structure provides (1) a map of tasks or responsibilities reporting relationships and grouping, and (2) mechanisms for linking and coordinating organizational elements into a cohesive whole. The crucial issue is the understanding of how to facilitate vertical and horizontal linkages through organization design. Classical organization theorists stressed vertical design and relied on vertical linkages to provide integration. Today, more attention has been focused on the need for horizontal coordination to enable organizations to adapt to the complex and rapidly changing environment. This aspect is a key issue in the internal functioning of schools: the level of interaction and integration between administration, teachers, and students.

Environment

The concept of "environmental domain" refers to establishing ongoing linkages and monitoring of those external "publics" that impact the organization or institution. In the case of a school, these publics might be the regulating agencies, students,

parents, and employers who are key members of the environmental domain. This area is similar to the need for stakeholder involvement, as advocated in the critical inquiry methodology.

Goals and Effectiveness

Traditional approaches used output goals, resource acquisition, or internal health and efficiency as the criteria of effectiveness. Newer approaches stress a broad view of effectiveness that consider these multiple activities simultaneously. This approach has evolved to solve the problem of different measures being more relevant to different constituencies. This is of particular importance for vocational education program evaluation, where give too much weight is often given to job placement rates in the determination of program effectiveness. This manuscript advocates that more consideration be paid to other aspects of "effectiveness" such as contextual or organization cultural variables, student benefits, and overall impact on the community.

Innovation and Change

Research in organizational theory has demonstrated that mechanistic (bureaucratic) structures are oriented toward technological stability, whereas organic (participative) structures are more suited to innovation and change. This is not meant to imply that bureaucracies do not change; they are, however, more suited to change in the administrative domain. This point is relevant in education where issues such as curriculum change imply the need for a more participative structure, whereas general rule changes are more effectively implemented administratively.

Decision Making

Daft (1983) summarized research findings relative to decision making:

> Most organizational decisions are not made in a logical, rational manner. Most decisions are not a series of steps that begins with problem identification, then analysis of alternatives and finally implementation of a certain solution. Decision processes are characterized by conflict, coalition-building, politics, trial and error, and mistakes. Intuition and hunch are often the criteria for choice. The decision process is disorderly, and may even seem random. In a few cases the solution may actually drive the problem because someone likes an idea and tries to find a problem as an excuse to adopt it.

> Another important idea is that individuals make decisions, but organizational decisions are not made by a single individual. Organizational decision-making is a social process. (p. 372)

This passage speaks to the reality of participation in the decision-making process. Organizations solve big problems through a series of small steps. Daft (1983) also noted that "the greater amount of conflict and coalition-building occurs when goals are not understood and agreed upon" (p. 373). This supports the rationale for stakeholders' involvement, be they from the external environment or from within the school.

Power and Politics

Power and politics can be viewed along a continuum; on one side is the rational model where the management of organizations is basically an engineering problem. This view assumes that organizations have specific goals, and environment and technical problems can be identified and logically solved. At the other end of the continuum is a political model of organization. Here the goals of the organization are not specific or agreed upon. Persons in different departments have different values, interests, and different desires for organizational outcomes. Decisions are made on the basis of power and political influence. Bargaining, negotiation, persuasion, and coalition building decide outcomes. This reality of power and political processes in organizations is an important aspect in the operation of an educational institution.

Intergroup Relationships

Organizational research has changed its position on conflict from the human relations view that conflict is bad for the organization and should be eliminated to the pluralistic view that conflict is inevitable and can be channeled beneficially.

> An organization without conflict is not sufficiently differentiated in goals, skills, and attitudes to be successful. Conflict is a sign of health and energy, and should be controlled only so that it does not get out of hand. (Daft 1983, p. 434)

With this acceptance of conflict, negotiation approaches have moved from win-lose configurations to win-win strategies (Johnson and Johnson 1975). Again, this reality supports the concept of involvement and ongoing discourse advocated by the critical inquiry approach.

Information and Control

Research on the organizational requirements for information and control systems has found that organizational requirements are very different from information and control systems designed for machines. Machine-based theories do not apply to human organizations.

Daft (1983) has developed a matrix on the "amount" and "richness" of information which suggests that organization and information structures are key areas on which organizations focus attention. Information amount is the volume of data about organizational activities that are gathered and interpreted by organization participants. The higher the uncertainty in the environment and the greater the number of problems occurring within the organization, the greater the amount of information that has to be processed. Information richness pertains to the carrying capacity of data. Some data cues are extremely informative for recipients, whereas other cues provide little understanding. The richest form of information is face-to-face interaction because it conveys several information cues simultaneously, including the spoken message, body language, and facial expression. Face-to-face interaction also provides immediate feedback so that understanding can be checked and misinterpretations corrected. Meherabian (1971) found that 93 percent of meaning conveyed is by tone of voice and facial expression. The least rich information is conveyed through numerical reports. Finally, information richness is important because it relates to the analyzability of problems.

> When a manager's task or problem is unanalyzable, the factors surrounding it are unclear and poorly understood, the manager has to acquire significant understanding in order to respond appropriately. Rich media, such as face-to-face, provide a variety of cues simultaneously that add greatly to the manager's understanding. If the manager relies on low rich media, such as a written report or a numerical document, the problem or solution will be over-simplified. The intangible, emotional, and difficult elements associated with unanalyzable problems can only be conveyed through rich media. On the other hand, when a problem is simple, a low rich medium is perferable. A written statement allows the verbal content to be transmitted uncluttered by other cues. In a face-to-face meeting, non-verbal cues may disagree with the simple message and cause confusion and misunderstanding. (Daft 1983, p. 302)

If one looks at the issue of participation and stakeholder involvement in this light, then clearly there are times when involvement and participation are crucial or at other times

unnecessary. It is the understanding of this distinction that is of major importance to organizational or institutional effectiveness.

Top Management Domain

Top management research demonstrates that the quality of top management makes a difference. Mintzberg (1973) identified 10 roles as being important in interaction with other persons: figurehead, leader, liaison person, information processor and monitor, disseminator, spokesperson, entrepreneur, disturbance handler, resource allocator, and negotiator. Additionally, current research has demonstrated that top managers have an impact on the organization culture (Peters and Waterman 1982 and Moss-Kanter 1983).

Within the educational setting, these findings demonstrate the importance of the chief administrator; local, State, and Federal administrators; and policymakers in encouraging the appropriate climate for relevant participation and stakeholder involvement to take root and flourish.

These findings from the business world are appropriate to educational settings for two important reasons:

- Some of the current controversy raging in education between the qualitative and quantitative camps would be resolved by applying the findings which clearly demonstrate that for different types of problems and issues, different strategies are necessary.

- Organizational or institutional effectiveness is influenced by involving the stakeholder in the decision-making process. Although these efforts are multifaceted and complex, those institutions that are able to adapt and respond to information from various internal and external sources are more likely to be successful.

Critical inquiry methodology offers a process whereby the aspects necessary for program improvement and organizational effectiveness can be addressed.

Admittedly, this has been a brief and superficial discussion of highly complex subject matter. However, the discussion should be sufficient to suggest that vocational curriculum is more than mere technical content. The notion that vocational education curriculum should have normative and practical content in addition to technical content was proposed. Moreover, it was suggested that learning is more than a technological model of inputs and outputs. Although Bloom's model resembles a technological framework, it is a starting point for critical discussion of the

learning process. The discussion on effective schools also suggests that schooling is a complex process and that any efforts to evaluate vocational education should view vocational education in that whole process. In essence, this manuscript suggests that critical inquiry appears to be a fruitful alternative for the evaluation of vocational education programs.

REFERENCES

Abramson, T. "Issues and Models in Vocational Education Evaluation." In Handbook of Vocational Education Evaluation, edited by T. Abramson, C. K. Tittle, and G. Cohen. Beverly Hills: Sage Publications, 1979.

ABT Associates. Implementation of the Educational Amendments of 1976: A Study of State and Local Compliance and Evaluation Practices. Final report. Cambridge, MA: Abt Associates, December 1980.

Alkin, M. C., and Fitz-gibbon, C. T. "Methods and Theories of Evaluating Programs," Journal of Research and Development in Education 8 (1975): 2-15.

The American Heritage Dictionary of the English Language, edited by W. Morris. Boston: Houghton Mifflin, 1976.

Aoki, T., and Harrison, E. "A Critical Approach to Program Evaluation in the B. C. Social Studies Assessment." Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April 1979.

Apple, M. W. "Social Evaluation of Curriculum," Educational Evaluation and Policy Analysis 5, no. 4 (Winter 1983): 425-434.

Ary, D.; Jacobs, F. C.; and Razavieh, A. Introduction to Research in Education. New York: Holt, Rinehart and Winston, 1979.

Averich, H. A.; Carroll, S. J.; Donaldson, T. S.; Kiesling, H. J.; and Pincus, J. How Effective Is Schooling? A Critical Review and Synthesis of Research Findings. Santa Monica, CA: The Rand Corporation, 1972.

Baker, E. L. New Directions in Evaluation Research: Implications for Vocational Education. Occasional Paper no. 55. Columbus: The National Center for Research in Vocational Education, The Ohio State University, November 1979.

Becker, G. S. Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. New York: National Bureau of Economic Research, 1964.

Bernstein, R. J. The Restructuring of Social and Political Theory. Philadelphia: University of Philadelphia Press, 1978.

Bloom, B. S. Human Characteristics and School Learning. New York: McGraw-Hill Book Co., 1982.

Bobbitt, F. The Curriculum. New York: Houghton and Mifflin, 1918.

Bowers, C. A. "The Reproduction of Technological Consciousness: Locating the Ideological Foundations of a Radical Pedagogy." Teachers College Record 83, no. 4 (Summer 1982): 529-557.

_____. The Promise of Theory: Education and the Politics of Cultural Change. New York: Longman, 1984.

Bowles, S. S., and Gintis, J. Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life. New York: Basic Books, 1976.

Bredo, E., and Feinberg, W. Knowledge and Values in Social and Educational Research. Philadelphia: Temple University Press, 1982.

Campbell, D. T., and Stanley, J. C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand-McNally Publishing Co., 1963.

Carnevale, A. P. Human Capital: A High Yield Corporate Investment. National Issues Series. Washington, DC: American Society for Training and Development, 1983.

Carroll, J. B. "A Model of School Learning." Teachers College Record 64 (1963): 722-733.

Chamberlaine, N. W.; Cullen, D. E.; and Lewis, D. "The Labor Market." In The Labor Sector. New York: McGraw-Hill Book Co., 1980.

Cherryholmes, C. H. "Discourse and Criticism in the Social Studies Classroom." Theory and Research in Social Education 9, no. 4 (Winter 1982): 57-73.

_____. "Knowledge, Power and Discourse in Social Studies Education." Journal of Education 165, no. 4 (1983): 341-358.

Cicirelli, V. G.; Evans, J. W.; and Schiller, J. S. "The Impact of Head Start: A Reply to the Report Analysis." Harvard Educational Review 40 (1970): 105-129.

Cobb, B. R., and Preskill, H. "Evaluating Vocational Education: A Review of Past Practices and Suggestions for Future Policy." Educational Evaluation and Policy Analysis 5, no. 4 (Winter 1983): 495-505.

Coleman, J. S.; Campbell, E. G.; Hobson, C. J.; McPartland, J. M.; Mood, A. M.; Weinfeld, F. D.; and York, R. L. Equality of Educational Opportunity. Washington, DC: Government Printing Office, 1966.

Committee on Vocational Education Research and Development. Assessing Vocational Education Research and Development. Washington, DC: National Academy of Sciences, 1976.

Cooley, W. W., and Leinhardt. G. "The Instructional Dimensions Study." Educational Evaluation and Policy Analysis 2, no. 1 (1980): 7-25.

Coomer, D. J. "A Critical Study of Educational Evaluation Theory and Practice." Ph.D. diss., The University of Minnesota, 1981.

Copa, G. H. "The Economics of Planning Vocational Education." In Vocational Education and the Nation's Economy, 7th Yearbook of the American Vocational Association, edited by W. G. Meyer. Washington, DC: American Vocational Association, 1977.

Cumming, R. D. Human Nature and History. Chicago: University of Chicago Press, 1969.

Daft, Richard. Organization Theory and Design. St. Paul, MN: West Publishing Co., 1983.

Darcy, R. Vocational Education Outcomes: Perspective for Evaluation. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1979.

Denham, C., and Lieberman, A., eds. Time to Learn. Washington, DC: Program on Teaching and Learning, National Institute of Education, 1980.

Desy, J.; Campbell, P. B.; and Gardner, J. A. High School Vocational Education Experiences: In School and in the Labor Market. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1984.

Dryzek, J. "Policy Analysis as a Hermeneutic Activity." Policy Sciences 14 (1982): 309-329.

Eisner, E. W. "Educational Connoisseurship and Criticism: Their Form and Functions in Educational Evaluation." In Evaluation Models: Viewpoints on Educational and Human Services Evaluation, edited by G. F. Madaus, M. S. Scriven, and D. J. Stufflebeam. Boston: Kluwer-Nijhoff Publishers, 1983.

_____. "The Use of Qualitative Forms of Evaluation for Improving Educational Practice." Educational Evaluation and Policy Analysis 6 (1979): 1-19.

Evans, R. N., and Herr, T. Foundations of Vocational Education 2d ed. Columbus: Charles E. Merrill Publishing Co., 1978.

Feinberg, W. Understanding Education. Cambridge: Cambridge University Press, 1983.

Freire, P. Pedagogy of the Oppressed. 2d ed. New York: Continuum, 1981.

Gallinelli, J. "Vocational Education Programs at the Secondary Level: A Review of Development and Purpose." In Handbook of Vocational Education Evaluation, edited by T. Abramson, C. K. Tittle, and L. Cohen. Beverly Hills: Sage Publications, 1979.

Gilsinan, J. F., and Volpe, L. C. "Do Not Cry Wolf until You Are Sure: The Manufactured Crises in Evaluation Research." Policy Sciences 17 (1984): 179-191.

Goffman, E. Asylums: Essays on the Social Situation of Mental Patients and Other Inmates. New York: Doubleday Anchor Books, 1961.

Goodlad, J. I. A Place Called School. New York: McGraw-Hill Book Co., 1984.

Gouldner, A. "Sociology: Contradictions and Intrastructure." In Knowledge and Values in Social and Educational Research, edited by E. Bredo and W. Feinberg. Philadelphia: Temple University Press, 1982.

Grasso, J. T. Impact Evaluation in Vocational Education: The State of the Art. Information Series no. 157. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1979.

Guba, E., and Lincoln, Y. S. Effective Evaluation: Improving the Usefulness of Evaluation Results through Responsive and Naturalistic Approaches. San Francisco: Jossey-Bass Publishers, 1981.

Habermas, J. Communication and the Evolution of Society. Boston: Beacon Press, 1979.

_____. Knowledge and Human Interests. Boston: Beacon Press, 1971.

_____. Legitimation Crisis. Boston: Beacon Press, 1975.

_____. "A Postscript to Knowledge and Human Interests." Philosophy of Social Science 3 (1973): 157-189.

_____. Theory and Practice. Boston: Beacon Press, 1973.

_____. Toward a Rational Society. Boston: Beacon Press, 1970.

_____. "Towards a Theory of Communicative Competence." Inquiry 13 (1970): 360-375.

Halasz, I. M., and Behm, K. S. Time on Task in Selected Vocational Education Classes. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1983.

Hall, G. E., and Jones, H. L. Competency-Based Education: A Process for the Improvement of Education. Englewood Cliffs, NJ: Prentice-Hall, 1976.

Hanson, G. A., and Parker, E. C. "The Vocational Education Industry." In Vocational Education and the Nation's Economy, 7th yearbook of the American Vocational Association, edited by W. G. Meyer. Washington, DC: American Vocational Association, 1977.

Held, D. Introduction to Critical Theory: Horkheimer to Habermas. London: J. Hutchinson Co., 1980.

Hellmuth, J., ed. Compensatory Education: A National Debate. New York: Bruner/Mazel, 1970.

Hempel, C. G. "Science and Human Values." In Aspects of Scientific Explanation and Other Essays in the Philosophy of Science, edited by C. G. Hempel. New York: The Free Press, 1965.

Hendrickson, G. Evaluating Vocational Education: The Federal Stimulus. Vocational Education Study Publication no. 5. Washington, DC: Government Printing Office, 1981.

House, E. R. "Assumptions Underlying Evaluation Models." Educational Researcher 7 (March 1978): 4-12.

_____. Evaluating With Validity. Beverly Hills: Sage Publications, 1980.

_____. "Evaluation as Scientific Management in U.S. School Reform." Comparative Education Review 22, no. 3 (October 1978): 388-401.

_____. "How We Think About Evaluation." In Philosophy of Evaluation. New Directions for Program Evaluation. San Francisco: Jossey-Bass, 1983.

House, E. R.; Glass, G. V.; McLean, L. D.; and Walder, D. F. "No Simple Answer: Critique of the Follow Through Evaluation." Harvard Educational Review 48 (1978): 128-160.

Hu, T. W., and Stromsdorfer, E. W. "Cost-Benefit Analysis of Vocational Education." In Handbook of Vocational Education Evaluation, edited by T. Abramson, C. K. Tittle, and T. Cohen. Beverly Hills: Sage Publications, 1979.

Jackson, P. W. Life in Classrooms. New York: Holt, Rinehart and Winston, 1968.

Jay, M. The Dialectical Imagination: A History of the Frankfurt School and the Institute of Social Research 1923-50. Boston: Little, Brown and Co., 1973.

Jensen, A. R. "How Much Can We Boost IQ and Scholastic Achievement?" Harvard Educational Review 39 (1969): 1-24.

Johnson, D. W., and Johnson, Frank P. Joining Together: Group Theory and Group Skills. Englewood Cliffs, NJ: Prentice-Hall, 1975.

Kliebard, H. M. "Bureaucracy and Curriculum Theory." In Curriculum and Evaluation, edited by A. A. Bellack and H. M. Kliebard. Berkeley: McCutchan Publishing Corporation, 1977.

Kuhn, S. The Structure of Scientific Revolutions. 2d ed. Chicago: The University of Chicago Press, 1970.

Lakatos, I. "Falsification and the Methodology of Scientific Research Programmes." In Criticism and the Growth of Knowledge, edited by I. Lakatos and A. Musgrave. Cambridge, England: University of Cambridge Press, 1970.

Lakatos, I., and Musgrave, A., eds. Criticism and the Growth of Knowledge. Cambridge, England: University of Cambridge Press, 1970.

Larson, M. E. Review and Synthesis of Research: Analysis for Curriculum Development in Vocational Education. Washington, DC: Office of Education, 1969. (ERIC Document Reproduction Service No. ED 035 746)

Lauden, L. Progress and Its Problems. Berkeley: University of California Press, 1970.

Leakcock, E. B. Teaching and Learning in City Schools. New York: Basic Books, 1969.

Lemish, P. S. "The Technical Approach and the Praxis Approach to Curriculum Development." Paper presented at the annual meeting of the American Educational Research Association, New York, New York, March 1982.

Lewis, D. R. "An Introduction to the Economics of Education." In Vocational Education and the Nation's Economy. 7th Yearbook of the American Vocational Association, edited by W. G. Meyer. Washington, DC: American Vocational Association, 1977.

Louch, A. R. Explanation and Human Action. Berkeley: University of California Press, 1966.

Lowi, T. J. The End of Liberalism: Ideology, Policy, and the Crisis of Public Philosophy. New York: W. W. Norton and Co., 1969.

MacDonald, J. B. "Curriculum Theory." In Curriculum Theorizing: The Reconceptualists edited by W. Punar. Berkely: McCuthan Publishing, 1975.

MacIntyre, A. "The Idea of a Social Science." In Knowledge and Values in Social and Educational Research, edited by E. Bredo and W. Feinberg. Philadelphia: Temple University Press, 1982.

MacKenzie, D. E. "Research for School Improvement: An Appraisal of Some Recent Trends." Educational Researcher 12 (1983): 5-17.

Mathison, S. "Exploring Alternative Forms of Explanation in Program Evaluation: Explanation Through Critical Theory." Paper presented at the joint annual meeting of the Evaluation Network and the Evaluation Research Society, Baltimore, Maryland, September 1982.

McCarthy, T. A. "A Theory of Communicative Competence." Philosophy of the Social Sciences 3 (1973): 135-156.

McCutcheon, G. "What in the World Is Curriculum Theory?" Theory Into Practice 21, no. 1 (1982): 18-22.

Meherabian, A. Silent Message. Belmont, CA: Wadsworth, 1974.

Mertens, D. M.; McElwain, D.; Garcia, G.; and Whitmore, M. The Effects of Participating in Vocational Education: Summary of Studies Reported since 1968. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1980.

Mintzberg, Henry. The Nature of Managerial Work. New York: Harper and Row, 1983.

Moss, J., Jr. "Education or Training: What's in a Name." VocEd 58, no. 1 (January-February 1983): 20.

Moss, J., Jr., and Smith, B. B. eds. Report of a Seminar: Process and Techniques of Vocational Curriculum Development. Minneapolis: Minnesota Research Coordinating Unit for Vocational Education, University of Minnesota, April 1970.

Moss-Kanter, R. The Change Masters: Innovation for Productivity in the American Corporation. New York: Simon and Schuster, 1983.

National Institute of Education. The Vocational Education Study: The Final Report. Vocational Education Study, Publication no. 8. Washington, DC: U.S. Department of Education, September 1981.

Norton, R. E. "CBE: A Humanistic and Realistic Approach to Technical and Occupational Education for the '80s." Paper presented at the "Special Programs: Developing Human Potential" workshop, sponsored by the Office of Special Programs, State University of New York, April 1984.

Panel of Consultants on Vocational Education. Education for a Changing World of Work. Washington, DC: Government Printing Office, 1964.

Patton, M. Q. Alternative Evaluation Research Paradigms. Grand Forks, ND: Center for Teaching and Learning, The University of North Dakota, 1975.

_____. Qualitative Evaluation Methods. Beverly Hills: Sage Publications, 1980.

_____. Utilization Focused Evaluation. Beverly Hills: Sage, 1978.

Perrow, C. Organization Analysis: A Sociological Approach. Belmont, CA: Wadsworth, 1970.

Peters, T., and Waterman, R. In Search of Excellence: Lessons from America's Best Run Companies. New York: Harper and Row, 1982.

Phillips, D. C. "After the Wake: Post-Positivistic Educational Thought." Educational Researcher 12, no. 5 (May 1983): 4-12.

_____. Theories, Values and Education. Melbourne, Australia: Melbourne University Press, 1971.

_____. "What Do the Research and the Practitioner Have to Offer Each Other?" Educational Research 9, no. 11 (December 1980): 17-24.

Provus, M. N. Discrepancy Evaluation. Berkeley: McCutchan Publishing Corporation, 1971.

Purkey, S. C., and Smith, M. S. "Too Soon to Cheer? Synthesis of Research on Effective Schools." Educational Leadership 40, no. 3 (1982): 64-69.

Rivlin, A. M. Systematic Thinking for Social Action. Washington, DC: Brookings Institution, 1971.

Rossi, P. H.; Freeman, H. E.; and Wright, S. R. Evaluation: A Systematic Approach. Beverly Hills: Sage Publications, 1979.

Schuetz, A. "Concept and Theory Formation in the Social Sciences." Journal of Philosophy 51 (April 1954): 257-273.

Scriven, M. S. "Evaluation Ideologies." In Evaluation Models: Viewpoints on Educational and Human Services Evaluation, edited by G. F. Madaus, M. S. Scriven, and D. G. Stufflebeam. Boston: Kluwer-Nijhoff Publishers, 1983.

_____. Evaluation Thesaurus. 2d ed. Inverness, CA: Edgepress, 1980.

_____. "Goal-Free Evaluation." In School Evaluation: The Politics and Process edited by E. R. House. Berkeley: McCutchan Publishing, 1973.

_____. "The Methodology of Evaluation." In Curriculum and Evaluation, edited by A. A. Bellack and H. M. Kliebard. Berkeley: McCutchan Publishing, 1977.

Searle, J. Speech Acts. Cambridge, England: Cambridge University Press, 1969.

Selltiz, C.; Jahoda, M.; Deutsch, M.; and Cook, S. W. Research Methods in Social Relations. 2d ed. New York: Holt, Rinehart and Winston, 1959.

Silberman, H. F. Determining Goals for Vocational Education. Occasional Paper no. 96. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1983.

Sirotnik, K. A. "An Outcome-Free Conception of Schooling: Implications for School-Based Inquiry and Information Systems." Educational Evaluation and Policy Analysis 6, no. 3 (Fall 1984): 227-239.

_____. Principles and Practice of Contextual Appraisal for Schools. Occasional Paper No. 5. Los Angeles: Laboratory in School and Community Education, University of California, 1984.

Sirotnik, K. A., and Oakes, J. "A Contextual Appraisal System for Schools: Medicine or Madness?" Educational Leadership 39, no. 3 (December 1981): 164-173.

_____. Critical Inquiry and School Renewal: A Liberation of Method within a Critical Theoretical Perspective. Occasional Paper no. 4. Los Angeles: Graduate School of Education, The University of California, 1983.

Smith, B. B., and Moss, J., Jr. Process and Techniques of Vocational Curriculum Development. Minneapolis: Research Coordinating Unit for Vocational Education, University of Minnesota, April 1970.

Smith, J. K. "Quantitative Versus Interpretive: The Problem of Conducting Social Inquiry." In Philosophy of Evaluation: New Directions for Program Evaluation, edited by E. R. House. San Francisco: Jossey-Bass, 1983.

_____. "Quantitative Versus Qualitative Research: An Attempt to Clarify the Issue." Educational Researcher 12, no. 3 (March 1983): 6-13.

Smith, M. S., and Bissell, J. S. "Report Analysis: The Impact of Head Start." Harvard Educational Review 40 (1970): 41-104.

Stake, R. E. "The Case Study Method in Social Inquiry." Educational Researcher 7, no. 2 (February 1978): 5-8.

_____. "Program Evaluation, Particularly Responsive Evaluation." In Evaluation Models: Viewpoints on Educational and Human Services Evaluation, edited by G. F. Madaus, M. S. Scriven, and D. J. Stufflebeam. Boston: Kluwer-Nijhoff Publishers, 1983.

Stufflebeam, D. L. "The CIPP Model for Program Evaluation." In Evaluation Models: Viewpoints on Educational and Human Services Evaluation, edited by G. F. Madaus, M. S. Scriven, and D. G. Stufflebeam. Boston: Kluwer-Nijhoff Publishers, 1983.

Stufflebeam, D. L., and Webster, W. J. "An Analysis of Alternative Approaches to Evaluation." In Evaluation Models: Viewpoints on Educational and Human Services Evaluation, edited by G. F. Madaus, M. S. Scriven, and D. G. Stufflebeam. Boston: Kluwer-Nijhoff Publishers, 1983.

Taba, H. Curriculum Development: Theory and Practice. New York: Harcourt, Brace and World, 1962.

Taylor, C. Hegel. Cambridge, England: Cambridge University Press, 1975.

_____. "Interpretation and the Sciences of Man." In Knowledge and Values in Social and Educational Research, edited by E. Bredo and W. Feinberg. Philadelphia: Temple University Press, 1982.

_____. "Neutrality in Political Science." In Social Structures and Political Theory, edited by W. E. Connally and G. Gordon. Washington, DC: Heath and Co., 1974.

Thompson, J. D. Organizations in Action. New York: McGraw-Hill, 1967.

Toulmin, S. Foresight and Understanding: An Enquiry into the Aims of Science. New York: Harper and Row Publishers, 1961.

Turnbull, B. J. "The Federal Role in Educational Improvement," Harvard Educational Review 52, no. 4 (November 1982): 57-71.

Tyler, R. W. Basic Principles of Curriculum and Instruction. Chicago: University of Chicago Press, 1950.

Violas, P. C. "Reflections on Theories of Human Capital, Skills Training, and Vocational Education." Educational Theory 31, no. 2 (Spring 1981): 137-151.

Walker, D. "Curriculum Theory Is Many Things to Many People." Theory into Practice 21, no. 1 (1982): 62-65.

Weick, K. E. "Educational Organizations and Loosely Coupled Systems." Administrative Science Quarterly 21 (1976): 1-19.

Weiss, C. H. "The Stakeholder Approach to Evaluation: Origins and Promise." In Stakeholder Based Education: New Directions for Program Evaluation, edited by A. S. Bryk. San Francisco: Jossey-Bass, 1983.

_____. "Utilization of Evaluation: Toward Comparative Study." In Evaluating Action Programs, edited by C. H. Weiss. Boston: Allyn and Bacon, 1972.

Winch, Peter. The Idea of a Social Science. New York: Humanities Press, 1958.

Wirth, A. G. Productive Work--In Industry and Schools: Becoming Persons Again. New York: University Press of America, 1983.

Wise, A. E. Legislated Learning: The Bureaucratization of the American Classroom. Berkeley: University of California Press, 1979.

Wittengenstein, J. Tractatus Logico-Philosphicus. 2d ed. Routledge and Kegan Paul, 1971.

Wolf, R. J. "Trial by Jury: The Process." Phi Delta Kappan 57 (November 1974): 185-187.

_____. "The Use of Judicial Evaluation Methods in the Formulation of Educational Policy." Educational Evaluation and Policy Analysis 1, no. 3 (May-June 1979): 19-28.

Woodward, J. Industrial Organization: Theory and Practice. London: Oxford University Press, 1965.

Worthington, R. M. Vocational Education in the United States: Retrospect and Prospect. Occasional Paper no. 101. Columbus: The National Center for Research in vocational Education, The Ohio State University, 1984.

Zais, R. S. Curriculum: Principles and Foundations. New York: Thomas Y. Crowell Company, 1976.

*U.S. GOVERNMENT PRINTING OFFICE: 1985/559-116/11523